CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Functional Analysis 220 (2005) 118–156
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

The short-cut test

Volker Metz*,[1]

Institut für Mathematik C, Technische Universität Graz, 8010 Graz, Austria

Received 13 June 2003; received in revised form 11 May 2004; accepted 7 June 2004
Communicated by L. Gross
Available online 23 November 2004

Abstract

The short-cut test detects existence and uniqueness of "Laplacians" on finitely ramified, graph-directed fractals. Previous results by Sabot, Nussbaum and the author are improved and extended. It opens up the way for further studies because it combines well established spectral, dynamical and analytic techniques. Its algorithmic and recursive structure is designed to provide computable and flexible criteria.
© 2004 Elsevier Inc. All rights reserved.

MSC: 31C25; 60J45; 47J10; 28A80

Keywords: Laplace operator; Diffusion; Nonlinear eigenvalue problem; Fractal

1. Introduction and results

Fractal sets are physically relevant because they resemble porous media. On them one would like to study diffusions or prototypically heat conduction. To construct it one has to find a "Laplace operator" Δ on the fractal F which then defines a heat semigroup via $P_t = e^{\Delta t}$. We will equivalently look for the Dirichlet form $\mathcal{D}(f, g) := \langle -\Delta f, g\rangle_{L^2(F,\mu)}$ on the fractal equipped with its normalized Hausdorff measure μ. The existence and uniqueness of such Dirichlet forms, under additional symmetry, regularity and scaling assumptions, on so-called finitely ramified, graph-directed fractals is the

* Fax: +49-521-106-6480.
E-mail address: metz@mathematik.uni-bielefeld.dc (V. Metz).
[1] Supported by the START-project Y96-MAT of the Austrian Science Fund and the DFG research group "Spektrale Analysis, asymptotische Verteilungen und stochastische Dynamik".

0022-1236/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2004.06.008

topic of this article. “Finitely ramified” roughly means that one can ruin the connection of F by removing only finitely many points. “Graph-directed” allows us to mix a finite number of building blocks in the construction of the fractal. Unlike Lindstrøm’s existence results for the relatively small class of “nested fractals” no general existence or uniqueness result is proven. Instead a tool box, the short-cut test, is presented which should enable the reader to answer the existence or uniqueness question in a flexible algorithmic way. In the examples tested so far it was a real short cut as compared to previously available techniques. Especially, Sabot’s results in [32] on existence and uniqueness are improved. They are in turn an improvement of Barlow’s considerations in [3].

The article covers a wide class of new examples, because previous articles mainly dealt with constructions based on a single building block. Only two exceptions came to the knowledge of the author, the diamond and the Hany fractal [14,24]. Fractals which are not finitely ramified (infinitely ramified), like the Sierpinski carpet studied in [4], are not covered by this article.

On finitely ramified fractals the existence and uniqueness of a Dirichlet form $\mathcal{E}$ is known to be equivalent to the existence and uniqueness of a discrete self-similar Dirichlet form on a finite skeleton of the fractal, as described by Lindstrøm. Such a Dirichlet form is the eigenvector, located in the interior of a cone P, of a nonlinear renormalization map Λ. Up to nonlinearity this is the set up of classical Perron–Frobenius theory. We will use a nonlinear version of it known as Hilbert’s projective metric on cones. The geometric aspect of this theory is a “weakly” Gromov hyperbolic metric space with Hilbert’s metric h on P in the sense of [12], the spectral side is an interval calculus imitating classical results on the spectral radius due to Collatz, Wielandt and Nussbaum, and the dynamical side is the iteration of the h-nonexpansive map Λ also studied by Nussbaum. The nonexpansiveness implies an aggregated action of Λ on so called parts. It mimics the closed classes of states of a Markov chain and is a way to analyze the “irreducibility” of Λ. The aggregated action decomposes the original eigenvalue problem into a small collection of lower dimensional subproblems of the same type. This is the “cut” aspect of the short-cut test. On irreducible subproblems we have to rely on nonlinearity. An efficient way to study the nonlinear aspects of the dynamics of Λ are Gâteaux derivatives at the boundary of P. This leads to monotone convergence problems in which infinite “conductances” appear. This “short circuiting of electrical networks” is the “shorting” aspect of the short-cut test. Several of the above techniques only use qualitative properties of Λ and are, therefore, also of general interest.

Surprisingly, at least to the author, these techniques turned out to be closely related to Sabot’s arguments in [32]. As compared to his results the main differences are: graph-directed instead of single block fractals are considered, his very restrictive Assumption H is completely removed, “existence without uniqueness” is no more a blind spot, and further tests can be developed because well established techniques are used. Technically speaking, Proposition 24(ii) allows us to avoid Sabot’s Assumption H. This statement in turn is a consequence of the geometric Lemma 6 and dynamic Proposition 15. The present article suggests four tests, the function test, Sabot’s test, the eigenvalue test and Nussbaum’s test. They can be freely combined in a recursive and algorithmic framework. This flexibility is responsible for the practical relevance of the short-cut test. It is necessary because no single test covers all examples. Section 7 gives advises how to use the tests efficiently.

The organization of the article is as follows. Section 2 describes the set up and sketches Lindstrøm's reduction of the existence and uniqueness of $\mathcal{E}$ to a finite-dimensional eigenvalue problem for the renormalization map Λ. Section 3 is about nonlinear Perron–Frobenius theory. It explains why Λ does not expand Hilbert's distance h, sketches the hyperbolic geometry of h in Section 3.1, relates spectral properties of Λ to a specific interval calculus in Section 3.2 and describes the dynamics of the iterates of Λ in Section 3.3. Section 4 provides some monotone convergence results for Λ which are used in Section 5 to calculate Gâteaux derivatives of Λ at the boundary of P. The main results are obtained in Sections 6 and 7. Proposition 24 gives criteria in (20) under which a fixed point of $\tilde{\Lambda}$, a normalized version of Λ, at the boundary of P is nonattracting. Four tests are suggested which verify (20) in different ways. Section 7 studies nonexistence and reduces the total number of tests that have to be performed further. Corollary 32 might even be of general interest. Finally, Section 8 gives illustrating examples.

2. Lindstrøm's renormalization approach

The first three construction steps of our guiding example, the Hany fractal, are indicated in Fig. 1. The fractal is named after the two authors of [9] where it was first considered. We interpret the Hany fractal as a graph-directed construction in the sense of Mauldin and Williams [19,7]. The initial triangle in the first column of Fig. 1 is a subset of $\mathrm{R}^2_t := \mathrm{R}^2 \times \{t\}$ and the initial square of the same column a subset of $\mathrm{R}^2_s := \mathrm{R}^2 \times \{s\}$. To construct the refined version in the second column of Fig. 1 inside the initial versions we need 12 similitudes with contraction factor $\frac{1}{3}$: Let us say $\psi_1, \psi_2, \psi_3 : \mathrm{R}^2_t \to \mathrm{R}^2_t$ for the three reduced copies of the triangle inside the initial triangle, $\psi_4 : \mathrm{R}^2_s \to \mathrm{R}^2_t$ for the reduced square inside the initial triangle, $\psi_5, \ldots, \psi_8 : \mathrm{R}^2_t \to \mathrm{R}^2_s$ for the reduced triangles inside the initial square, and finally $\psi_9, \ldots, \psi_{12} : \mathrm{R}^2_s \to \mathrm{R}^2_s$ for the reduced squares inside the initial square. The graph directing the construction of the fractal has the vertices $V := \{s, t\}$ and the directed edges $E := \{1, \ldots, 12\}$, where $e \in E$ points from $t(e) := v$ to $d(e) := w$ when ψ_e has the domain R^2_w and the target R^2_v, for

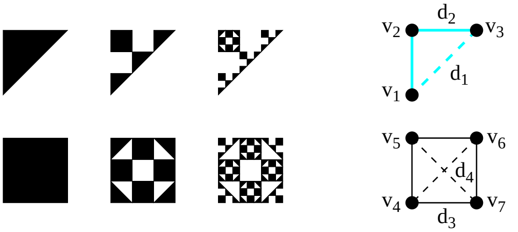

Fig. 1. The first three construction steps of the Hany fractal and the skeleton graph Γ_0 of the initial stage.

$v, w \in V$. According to [7, Theorem 4.3.5] there exist nonvoid compacts $K_v \subset \mathbb{R}^2_v$, $v \in V$, such that

$$K_v = \bigcup_{e \in t^{-1}(v)} \psi_e(K_{d(e)}).$$

The Hany fractal is $K_t \cup K_s \subset \mathbb{R}^2 \times \{s, t\}$. For $M \subset \mathbb{R}^2 \times \{s, t\}$ define the refinement map Ψ by

$$\Psi(M) := \bigcup_{e \in E} \psi_e(M \cap \mathbb{R}^2_{d(e)}).$$

Then F is a fixed point of Ψ and therefore called self-similar. The Hausdorff dimension of F is $\frac{\ln(7+\sqrt{17})-\ln 2}{\ln 3} \sim 1.56$ and the corresponding Hausdorff measure μ can be normalized to a probability measure [7, Theorem 6.4.8].

Each initial shape of the Hany fractal corresponds to a weighted graph which will be interpreted as an electrical resistor network (see [6]). Its vertices (of generation 0) are $\{v_1, \dots, v_7\} =: V_0$ placed at the corner points of the triangle and the square as indicated in the fourth column of Fig. 1. Define a conductance to be a function $c_0 : V_0^2 \to \mathbb{R}_+$ which is symmetric, vanishes on all edges not indicated in the graph Γ_0 of Fig. 1, named the skeleton graph for short, and has four real values $d_1, \dots, d_4 \geqslant 0$ on the colored edges as indicated in the figure. The initial triangle and its refinements are invariant under certain reflections which we extend by the identity to $\mathbb{R}^2_s$. Analogously the square and its refinements are invariant under further reflections which we extend by the identity on $\mathbb{R}^2_t$. By composition these extended reflections generate a symmetry group $\mathfrak{G}$ on $\mathbb{R}^2 \times \{s, t\}$. The conductance c_0 is $\mathfrak{G}$-invariant because of our arrangement of $d_1, \dots, d_4$, that is, $c_0(x, y) = c_0(\mathfrak{g}(x), \mathfrak{g}(y))$ for all $\mathfrak{g} \in \mathfrak{G}$ and $x, y \in V_0$. For $f, g : V_0 \to \mathbb{R}$ and $v \in V$ define discrete symmetric Dirichlet forms

$$\mathcal{E}_0^v(f, g) := \frac{1}{2} \sum_{x, y \in V_0 \cap K_v} (f(y) - f(x))(g(y) - g(x))c_0(x, y) \tag{1}$$

in the sense of [8]. Then $\mathcal{E}_0 := \mathcal{E}_0^t + \mathcal{E}_0^s$ is a Dirichlet form on V_0 which is $\mathfrak{G}$-invariant, that is, $\mathcal{E}_0(f, g) = \mathcal{E}_0(f \circ \mathfrak{g}, g \circ \mathfrak{g})$ for all $\mathfrak{g} \in \mathfrak{G}$ and $f, g : V_0 \to \mathbb{R}$. The Dirichlet operator E_0 of $\mathcal{E}_0$ is given by $\mathcal{E}_0(f, g) = \langle -E_0 f, g \rangle$, for $f, g : V_0 \to \mathbb{R}$, with respect to the euclidean inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{V_0}$.

Each symmetric Dirichlet form $\mathcal{E}$ with conductance $c_{\mathcal{E}}$ on a finite vertex set T defines a graph $\Gamma(\mathcal{E})$ with vertices T and edges $\{\{x, y\} \subset T \mid c_{\mathcal{E}}(x, y) > 0\}$. Whenever we use graph theoretic notions with respect to $\mathcal{E}$ we refer of course to $\Gamma(\mathcal{E})$. In our example every Dirichlet form $\mathcal{E}$ with $d_1, \dots, d_4 > 0$ has the property $\Gamma(\mathcal{E}) = \Gamma_0$. The latter has two connected components, the triangle $\{v_1, v_2, v_3\}$ and the square $\{v_4, \dots, v_7\}$.

For $n \in \mathbb{N}$ the vertices of generation n are given by $V_n := \Psi^n(V_0)$. Similarly we will define a Dirichlet form $\mathcal{E}_1$ on V_1. Because of polarization we can use quadratic

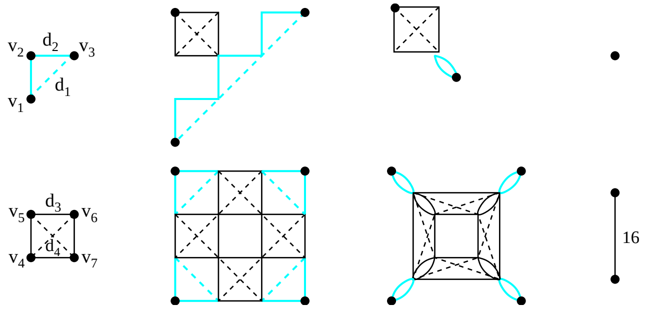

Fig. 2. The initial and the first generation graph of the Hany fractal and two short circuited versions of the latter.

forms, or energies for short, instead of Dirichlet forms. We set $\mathcal{E}(f) := \mathcal{E}(f, f)$ for any Dirichlet form $\mathcal{E}$ and any function f in its domain. Choose $\mathfrak{G}$-invariant coupling weights $\eta_e > 0$, $e \in E$, that is, whenever $\mathfrak{g} \circ \psi_e(K_{d(e)}) = \psi_{\tilde{e}}(K_{d(\tilde{e})})$ for some $\mathfrak{g} \in \mathfrak{G}$ and some $e, \tilde{e} \in E$, then we require $\eta_e = \eta_{\tilde{e}}$. In our guiding example we choose all coupling weights equal to 1. We define the coupling map Ψ for $f : V_0 \to \mathrm{R}$ by

$$\Psi(\mathcal{E}_0)(f) := \mathcal{E}_1(f) := \sum_{e \in E} \eta_e \cdot \mathcal{E}_0^{d(e)}(f \circ \psi_e).$$

Again $\mathcal{E}_1$ is $\mathfrak{G}$-invariant. The electrical resistor network defined by $\mathcal{E}_1$ on the Hany fractal is depicted in the second column of Fig. 2.

Next, we define the trace $\mathrm{Tr}(\mathcal{E}_1) := \mathrm{Tr}_{V_0}(\mathcal{E}_1)$ of $\mathcal{E}_1$ on V_0 in the sense of [8, Section 6] for $f : V_0 \to \mathrm{R}$ by

$$\mathrm{Tr}(\mathcal{E}_1)(f) := \inf\{\mathcal{E}_1(g) \mid g : V_1 \to \mathrm{R}, g\,|_{V_0} = f\}.$$

According to the Dirichlet principle, [13, Theorem 2.1.6], a minimizing element h solves the Dirichlet problem on the "open" set $V_1 \setminus V_0$ with "boundary" data f on V_0, in other words, h is $\mathcal{E}_1$-harmonic on $V_1 \setminus V_0$. Since $\mathcal{E}_1$ was $\mathfrak{G}$-invariant, its trace also is.

Finally, Lindstrøm's renormalization map Λ is defined to be

$$\Lambda := \mathrm{Tr} \circ \Psi.$$

It acts on the set D of all Dirichlet forms defined by $\mathfrak{G}$-invariant conductances as in (1). Let us denote the characteristic function of $M \subset F$ by 1_M and abbreviate $1_{\{x\}}$ to 1_x for $x \in F$. By $-\mathcal{E}_0(1_x, 1_y) =: c_{\mathcal{E}_0}(x, y)$, for all $x, y \in V_0$, every $\mathcal{E}_0 \in \mathrm{D}$ defines a $\mathfrak{G}$-invariant conductance. In our example the corresponding bijective map between the cone D and the cone R_+^4 of all $\mathfrak{G}$-invariant conductances on V_0 extends to a vector space isomorphism between $\mathrm{B} := \mathrm{D} - \mathrm{D}$ and R^4. We endow B with the sup norm

$\|\mathcal{E}\| := \sup\{|\mathcal{E}(f)|;\ f : V_0 \to \mathrm{R}, \|f\|_2 = 1\}$, where $\|\cdot\|_2$ is the euclidean norm on R^{V_0}. With respect to the $\|\cdot\|$-topology we denote the interior by $(\cdot)^\circ$, the boundary by $\partial(\cdot)$ and the closure by $\overline{(\cdot)}$. We will slightly abuse terminology and call an $\mathcal{E} \in \mathrm{D}$ irreducible when $\Gamma(\mathcal{E}) = \Gamma_0$ (more precisely, one should call it minimally reducible) and reducible otherwise (or not minimally reducible).

In B we have the cone $\mathrm{P} := \{\mathcal{E} \in \mathrm{B} | \mathcal{E}(\cdot) \geqslant 0\}$ of positive semidefinite forms. For $\mathcal{E} \in \mathrm{B}$ we denote $\{f : V_0 \to \mathrm{R} | \mathcal{E}(f) = 0\}$ by $\ker \mathcal{E}$. Every $\mathcal{E} \in \mathrm{P}^\circ$ has by definition a minimal kernel. Since $\mathrm{D}^\circ \subset \mathrm{P}^\circ$, the minimum principle, [13, Theorem 3.2.5], shows that this kernel consists of all functions constant on the triangle $\{v_1, v_2, v_3\}$ and on the square $\{v_3, \ldots, v_7\}$. The set of all irreducible Dirichlet forms is thus $\mathrm{D} \cap \mathrm{P}^\circ$. We want to solve the eigenvalue problem

$$\Lambda(\mathcal{E}) = \gamma\mathcal{E} \quad \text{for some } \mathcal{E} \in \mathrm{D} \cap \mathrm{P}^\circ \tag{2}$$

because of Lindtsrøm's reduction theorem below. The coupling map is linear but the trace map is not by Proposition 2(ii).

Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a conservative Dirichlet form in $L^2(F, \mu)$. We are interested in its space-energy scaling properties, that is, we look at all functions $f \in \mathcal{D}(\mathcal{E})$ which are constant on all but one component $K_v, v \in V$, for all $e \in E$ with $d(e) = v$ and ask whether $\mathcal{E}(f)$ is a constant multiple of $\mathcal{E}(f \circ \psi_e)$ which is independent of *f*. When the answer is "yes" for all components of *F*, then the Dirichlet form is termed self-similar. This corresponds to the scaling properties of the classical Dirichlet integral, the prototype of all Dirichlet forms. In the fractal case only the adapted scaling by a ψ_e is allowed because otherwise the image of the contraction might not be a subset of *F* anymore.

Theorem 1 (*Reduction, Lindstrøm [17]*). *The existence and uniqueness* (*up to positive multiples*) *of a local, regular, conservative, symmetric and self-similar Dirichlet form on $L^2(F, \mu)$ for which points have positive capacity is equivalent to the existence and uniqueness* (*up to positive multiples*) *of a solution to* (2) *for which each coupling weight is strictly bigger than γ.*

Proof (*Sketch*). "$\Leftarrow$" An analytic version of Lindstrøm's ideas can be found in [13, Chapter 3] for p.c.f. self-similar sets and in [9, Section 4] for finitely ramified, graph directed fractals. The basic idea is to use the eigenvalue in (2) as a scaling factor in $\mathcal{E}_n := \gamma^{-n}\Psi^n(\mathcal{E})$, for $n \in \mathrm{N}$. Thanks to the finite number of ramification points we are now in a situation similar to the discretization of the one-dimensional Laplacian on the unit interval: The approximating models $(\mathcal{E}_n)_n$ give exactly the values of the limiting model on the "grid points" $(V_n)_n$. The condition on the coupling weights ensures that all points will have positive capacity.

"$\Rightarrow$" Since points have positive capacity, the Dirichlet form $(\mathcal{D}, D(\mathcal{D}))$ on $L^2(F, \mu)$ has traces on V_0 and V_1. The self-similarity of $\mathcal{D}$ now means that there exist coupling weights $\{\eta_e | e \in E\}$ such that the multiples in the definition of "self-similar" are given by $\{\gamma \cdot \eta_e | e \in E\}$. $\square$

3. Nonlinear Perron–Frobenius theory

We will study Λ and see that it is P-concave and positively homogeneous. Thus the behavior of Λ is particularly close to the one of nonnegative matrices in classical Perron–Frobenius theory [15]. Especially, we can analyze our eigenvalue problem (2) by spectral and dynamical means. But in our case irreducibility is of limited use. Instead we have to rely on the nonlinearity.

For every $\mathcal{A} \in \mathrm{P}$ there exists a solution to the Dirichlet problem because the proof of the Dirichlet principle only uses bilinearity and positive semidefiniteness. When the solution is not unique we single out a unique one in requiring a minimal L^2-norm. For the "open" set $V_1 \setminus V_0$, the "boundary" data $f : V_0 \to \mathrm{R}$ and the positive semidefinite form $\mathcal{A}_1$ let us denote this unique solution by $H^{\mathcal{A}_1}_{V_1 \setminus V_0} f$. This is consistent with the probabilistic interpretation of harmonic functions and will be very useful in various continuity questions.

Proposition 2. (i) $\Lambda : \mathrm{D} \to \mathrm{D},\ \mathrm{P} \to \mathrm{P}$.

(ii) *For all* $\mathcal{A}, \mathcal{B} \in \mathrm{P},\ f : V_0 \to \mathrm{R}$ *and* $\alpha \geqslant 0$,

$$\Lambda(\alpha\mathcal{A}) = \alpha\Lambda(\mathcal{A}),$$

$$\Lambda(\mathcal{A}+\mathcal{B})(f) \geqslant \Lambda(\mathcal{A})(f) + \Lambda(\mathcal{B})(f).$$

Equality holds in the inequality if and only if $H^{\mathcal{A}_1}_{V_1 \setminus V_0} f - H^{\mathcal{B}_1}_{V_1 \setminus V_0} f$ *is an element of* $\ker(\mathcal{A}+\mathcal{B})_1$.

(iii) $\Lambda : \mathrm{D} \to \mathrm{D}$ *is continuous. Furthermore,* $\Lambda \in C^1(\mathrm{P}^\circ)$ *and the* (*Fréchet*) *derivative of* Λ *at* $\mathcal{A} \in \mathrm{P}^\circ$ *in direction* $\mathcal{D} \in \mathrm{B}$ *evaluated at* $f : V_0 \to \mathrm{R}$ *is* $D\Lambda_{\mathcal{A}}(\mathcal{D})(f) := \mathcal{D}_1(H^{\mathcal{A}_1}_{V_1 \setminus V_0} f)$.

(iv) *For* $\mathcal{A} \in \mathrm{P}^\circ$ *denote the operators of* $\Lambda(\mathcal{A})$ *and* $\mathcal{A}_1$ *by* $\Lambda(A)$ *and* A_1, *respectively. Define the "boundary"* $B := V_0$, *the "interior"* $I := V_1 \setminus V_0$ *and denote the submatrix of* A_1 *with lines L and rows R by* A_{LR}, *for* $L, R \in \{B, I\}$. *Then the Schur complement formula holds*:

$$\Lambda(A) = A_{BB} - A_{BI} A_{II}^{-1} A_{IB}.$$

Proof (*Sketch*). (i): $\Lambda : \mathrm{P} \to \mathrm{P}$ is a direct consequence of the definition, because the Dirichlet principle holds for positive semidefinite operators. Alternatively see [2, Theorem 3]. To prove $\Lambda : \mathrm{D} \to \mathrm{D}$ we have to check whether all "conductances" of $\Lambda(\mathcal{A})$ are nonnegative. This is a consequence of the minimum principle, which is valid because of the Markov property of Dirichlet forms. For details see [22, Proposition 2.1(a)].

(ii): The positive homogeneity of Λ follows from the definition. The superadditivity is a consequence of the Dirichlet principle [2, Theorems 5,9]. Hence there can be only additivity when $H^{\mathcal{A}_1}_{V_1 \setminus V_0} f$ is also $\mathcal{B}_1$-harmonic on $V_1 \setminus V_0$. Symmetrizing this statement we arrive at the desired assertion.

(iii): The continuity of Λ is a consequence of the minimum principle and our choice of harmonic functions with minimal L^2-norm [22, Proposition 2.1(b)]. The differentiability is [2, Theorem 2]. It is a consequence of the differentiability of the inversion in (iv).

(iv): The Schur complement formula is [1, Formula 6, Theorem 6]. Schur's formula appears because the trace is merely a restriction on Green's functions. Thus we have to invert partitioned matrices. □

The Schur complement formula can be used to prove well-known results of electrical network theory like: the resistance of resistors in line or parallel, the star-triangle transformation, the irrelevance of dangling ends and loops etc. We will freely use these arguments in the sequel.

Proposition 2(ii) advises us to estimate quotients of forms instead of differences. The cone P defines a partial ordering on B via $\mathcal{A} \leqslant \mathcal{B}$ if $\mathcal{B} - \mathcal{A} \in \mathrm{P}$. This means $\mathcal{A}(f) \leqslant \mathcal{B}(f)$ for all $f : V_0 \to \mathrm{R}$. Now Proposition 2(ii) says that Λ is positively homogeneous and P-increasing. Thus one should estimate quotients of forms to use these properties efficiently. Let $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$. When $\ker \mathcal{A} \subset \ker \mathcal{B}$, then

$$\inf[\mathcal{A}/\mathcal{B}] := \sup\{\alpha > 0 | \alpha\mathcal{B} \leqslant \mathcal{A}\}$$

is a positive real. Whenever $\ker \mathcal{B} \subset \ker \mathcal{A}$, then

$$\sup[\mathcal{A}/\mathcal{B}] := \inf[\mathcal{B}/\mathcal{A}]^{-1} = \inf\{\alpha > 0 | \mathcal{A} \leqslant \alpha\mathcal{B}\}$$

is a positive real. In the case of $\ker \mathcal{A} = \ker \mathcal{B}$, the range of the quotient

$$\left[\frac{\mathcal{A}}{\mathcal{B}}\right] := \left\{ \frac{\mathcal{A}}{\mathcal{B}}(f) \,\middle|\, f : V_0 \to \mathrm{R},\ f \notin \ker \mathcal{A} \right\} = [\inf[\mathcal{A}/\mathcal{B}], \sup[\mathcal{A}/\mathcal{B}]]$$

is a bounded interval in the positive reals.

Proposition 3. *Let* $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$. *In the case of* $\ker \mathcal{A} \subset \ker \mathcal{B}$, *we have the inequality* $\inf[\Lambda(\mathcal{A})/\Lambda(\mathcal{B})] \geqslant \inf[\mathcal{A}/\mathcal{B}]$ *and in the case of* $\ker \mathcal{B} \subset \ker \mathcal{A}$, *the inequality* $\sup[\Lambda(\mathcal{A})/\Lambda(\mathcal{B})] \leqslant \sup[\mathcal{A}/\mathcal{B}]$. *When* $\ker \mathcal{A} = \ker \mathcal{B}$, *then*

$$\left[\frac{\Lambda(\mathcal{A})}{\Lambda(\mathcal{B})}\right] \subset \left[\frac{\mathcal{A}}{\mathcal{B}}\right]. \tag{3}$$

Proof. The proof could be done in two lines but we will give an extended version which will be useful later. Let $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$ with $\ker \mathcal{A} \subset \ker \mathcal{B}$. Then $\inf[\mathcal{A}/\mathcal{B}]$ is a positive real and $\mathcal{D} := \mathcal{A} - \inf[\mathcal{A}/\mathcal{B}]\mathcal{B} \in \partial \mathrm{P}$, since P is $\|\cdot\|$-closed. By Proposition 2(ii) there exists $\mathcal{R} \in \mathrm{P}$ such that

$$\Lambda(\mathcal{A}) = \Lambda(\inf[\mathcal{A}/\mathcal{B}]\mathcal{B} + \mathcal{D}) = \inf[\mathcal{A}/\mathcal{B}]\Lambda(\mathcal{B}) + \Lambda(\mathcal{D}) + \mathcal{R}. \tag{4}$$

Thus $\inf[\Lambda(\mathcal{A})/\Lambda(\mathcal{B})] \geqslant \inf[\mathcal{A}/\mathcal{B}]$. The inequality for the suprema follows analogously. The inclusion of intervals is an immediate consequence of the two inequalities we just verified. □

3.1. Geometric aspects

The kind of nonlinear Perron–Frobenius theory we are going to use is known as Hilbert's projective metric on cones [28]. It turns the set of all rays in P° into a hyperbolic metric space similar to the open unit disc equipped with the Klein distance. Its adaptation to Λ and P can be found in [21]. The task of this section is to describe those aspects of the hyperbolic geometry which will be needed later.

Coinciding kernels define an equivalence relation on $\mathrm{P} \setminus \{0\}$. We call its equivalence classes P-parts. The P-part of an $\mathcal{A} \in \mathrm{P} \setminus \{0\}$ is denoted by $\mathrm{P}_\mathcal{A}$. The meaning of $\ker \mathrm{P}_\mathcal{A}$ now simply is $\ker \mathcal{A}$. For $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$ with $\mathrm{P}_\mathcal{A} = \mathrm{P}_\mathcal{B}$ define Hilbert's projective metric h on P by

$$h(\mathcal{A}, \mathcal{B}) := \ln \frac{\sup[\mathcal{A}/\mathcal{B}]}{\inf[\mathcal{A}/\mathcal{B}]}.$$

It is a pseudo distance because $h(\alpha\mathcal{A}, \beta\mathcal{B}) = h(\mathcal{A}, \mathcal{B})$ for all $\alpha, \beta > 0$. It vanishes if and only if $\mathcal{A}$ is a positive multiple of $\mathcal{B}$. The function $\mathcal{B} \mapsto h(\mathcal{A}, \mathcal{B})$ tends to $+\infty$ when $\mathcal{B}$ tends to a form with $\ker \mathcal{B} \neq \ker \mathcal{A}$. Therefore, we define $h(\mathcal{A}, \mathcal{B}) := +\infty$ for $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$ which do not lie in the same P-part. Let S be the unit sphere in $(\mathrm{B}, \|\cdot\|)$ and P_0 a P-part. Since $\|\cdot\|$ is P-increasing, $(\mathrm{P}_0 \cap S, h)$ is a complete metric space [28, Theoerm 1.2]. On $\mathrm{P}_0 \cap S$, h and the $\|\cdot\|$-metric are locally equivalent and thus define the same topology [28, Formulas 1.21, 1.22]. Finally, h-balls $B_r^h(\mathcal{C}) := \{\mathcal{A} \in \mathrm{P} | h(\mathcal{A}, \mathcal{C}) \leqslant r\}$, with $\mathcal{C} \in \mathrm{P} \setminus \{0\}$ and $r > 0$, are convex (but not necessarily strictly convex) [28, Lemma 4.1].

Let us partially order the set of all P-parts in saying $\mathrm{P}_1 \leqslant \mathrm{P}_0$ when $\ker \mathrm{P}_0 \subset \ker \mathrm{P}_1$. The meaning of $\mathrm{P}_1 < \mathrm{P}_0$ then is that $\ker \mathrm{P}_0$ is a strict subset of $\ker \mathrm{P}_1$.

Corollary 4. *Let $\mathcal{A}, \mathcal{B} \in \mathrm{P}$ with coinciding kernels and $\Lambda(\mathcal{A}), \Lambda(\mathcal{B}) \neq 0$. Then $h(\Lambda(\mathcal{A}), \Lambda(\mathcal{B})) \leqslant h(\mathcal{A}, \mathcal{B})$. That is, Λ is h-nonexpansive. Moreover, Λ respects the partial ordering of* P*-parts.*

Proof. The statement about h is an immediate consequence of (3). The remaining assertion follows from the inequalities in Proposition 3. □

The Vicsek set with reduced symmetry requirements in Section 8.1 has linear independent eigenvectors in $\mathrm{D} \cap \mathrm{P}^\circ$. Thus a strict h-contraction is in general not possible.

Note that Hilbert's distance h_{D}, defined by the cone D instead of P, might be expanded by Λ [21].

We can use (4) to identify sources of h-contraction. It implies that h is strictly contracted when $\Lambda(\mathcal{D}) \in \mathrm{P}^\circ$. Hence $\Lambda(\mathrm{P} \setminus \{0\}) \subset \mathrm{P}^\circ$ would imply strict h-contraction. This is the classical positivity argument of Perron [33, Theorem 1.1]. We will see in Sections 3 and 6 that the Hany fractal has several eigenvectors in $\partial\mathrm{P}$. So positivity does not help in general. Again by (4) the Hilbert distance is also contracted by $\mathcal{R} \in \mathrm{P}^\circ$. Thus strict superadditivity is another source of h-contraction! The Gâteaux derivative techniques of Section 5 are designed to use nonlinearity efficiently.

Our last task is to compare the $\|\cdot\|$- and the h-distance near $\partial\mathrm{P}$. For $\mathcal{A} \in \mathrm{P}$ we have defined $\ker\mathcal{A}$. Let us denote the euclidean orthogonal complement of $\ker\mathcal{A}$ in R^{V_0} by $\mathrm{im}\,\mathcal{A}$, and again $\mathrm{im}\,\mathrm{P}_{\mathcal{A}} := \mathrm{im}\,\mathcal{A}$.

Lemma 5. *The Hilbert distance* $h : \mathrm{P} \setminus \{0\} \to \mathrm{R}_+ \cup \{+\infty\}$ *is lower semi-continuous on* $((\mathrm{P} \setminus \{0\})^2, \|\cdot\| + \|\cdot\|)$.

Proof. Let $(\mathcal{A}^n)_n, (\mathcal{B}^n)_n \subset \mathrm{P} \setminus \{0\}$ be two $\|\cdot\|$-convergent sequences with limits $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$, respectively. For every $f : V_0 \to \mathrm{R}$ the map $\mathcal{C} \mapsto \mathcal{C}(f)$ is $\|\cdot\|$-continuous on B. Thus there exists a $\|\cdot\|$-neighborhood V of $\mathcal{A}$ which consists only of $\mathcal{C} \in \mathrm{P}$ with $\ker\mathcal{C} \subset \ker\mathcal{A}$. Denote the analogous $\|\cdot\|$-neighborhood of $\mathcal{B}$ by W. Without loss of generality $(\mathcal{A}^n)_n \subset V$ and $(\mathcal{B}^n)_n \subset W$. When $\mathrm{P}_{\mathcal{A}} \neq \mathrm{P}_{\mathcal{B}}$, then there exists an $f \in \ker\mathcal{A} \setminus \ker\mathcal{B}$ or a $g \in \ker\mathcal{B} \setminus \ker\mathcal{A}$. Without loss of generality let us assume the first, that is, $\mathcal{A}(f) = 0 < \mathcal{B}(f)$. Then there exists another $\|\cdot\|$-neighborhood $W' \subset W$ of $\mathcal{B}$ such that $\mathcal{C} \mapsto \mathcal{C}(f)$ is uniformly bounded below on W'. Thus

$$\lim_{n\to\infty} \inf h(\mathcal{A}^n, \mathcal{B}^n) = +\infty = h(\mathcal{A}, \mathcal{B}).$$

Now let $\mathrm{P}_{\mathcal{A}} = \mathrm{P}_{\mathcal{B}}$. Since we are interested in the limes inferior of $h(\mathcal{A}_n, \mathcal{B}_n)$ we suppose without loss of generality that $h(\mathcal{A}_n, \mathcal{B}_n) < \infty$ for all $n \in \mathrm{N}$. So both sequences are contained in the same P-part P_0 and $\mathrm{P}_{\mathcal{A}} = \mathrm{P}_{\mathcal{B}} \subset \overline{\mathrm{P}_0}$. When $\mathrm{P}_0 = \mathrm{P}_{\mathcal{A}}$ then the $\|\cdot\|$- and the h-topology coincide. Hence h is $\|\cdot\|$-continuous. When $\mathrm{P}_{\mathcal{A}} < \mathrm{P}_0$ then there exists a $\|\cdot\|$-neighborhood W of the line segment from $\mathcal{A}$ to $\mathcal{B}$ consisting of forms whose kernels are subsets of $\ker\mathcal{A}$. Thus $[\mathcal{A}_n/\mathcal{B}_n]$ is calculated for a bigger set of functions than $[\mathcal{A}/\mathcal{B}]$. This implies the desired inequality. □

Let $\mathrm{P}_1 \subset \partial\mathrm{P}$ be a P-part and $\mathrm{P}_1 \subset U \subset \mathrm{B}$. We say that U is infinitely h-deep (with respect to P_1), when for every $r > 0$ the r-body of the h-distance

$$(\mathrm{P}^\circ \setminus U)_r := \bigcup_{\mathcal{A} \in \mathrm{P}^\circ \setminus U} B_r^h(\mathcal{A})$$

does not $\|\cdot\|$-accumulate at P_1.

Lemma 6. *Let* $\mathrm{P}_1 \subset \partial\mathrm{P}$ *be a* P*-part. Then every* $\|\cdot\|$*-neighborhood* U *of* P_1 *is infinitely* h*-deep.*

Proof. All topological notions of this proof are with respect to the $\|\cdot\|$-topology on B. Without loss of generality assume that U is open. Let $r > 0$ and suppose $(\mathrm{P}^\circ \setminus U)_r$ accumulates at $\mathcal{B} \in \mathrm{P}_1$. Then there exists a sequence $(\mathcal{C}^n)_n \subset \mathrm{P}^\circ \setminus U$ and a sequence $(\mathcal{P}^n)_n \subset \mathrm{P}$ converging to $\mathcal{B}$ such that $h(\mathcal{C}^n, \mathcal{B}^n) \leqslant r$ for all $n \in \mathrm{N}$. Since h is constant on rays, we can project everything to S. Now $\mathrm{P} \cap S$ is compact and $(\mathrm{P} \setminus U) \cap S$ is a closed subset. Thus the projected $(\mathcal{C}^n)_n$ has an accumulation point $\mathcal{C} \in \mathrm{P} \setminus U$ approximated by the projected subsequence $(\mathcal{C}^{n_k})_k$. By Lemma 5

$$h(\mathcal{C}, \mathcal{B}) \leqslant \lim_{k\to\infty} \inf\, h(\mathcal{C}^{n_k}, \mathcal{P}^{n_k}) \leqslant r < \infty.$$

But $\mathcal{C} \in (\mathrm{P} \setminus U) \subset (\mathrm{P} \setminus \mathrm{P}_1)$ and therefore $h(\mathcal{C}, \mathcal{B}) = \infty$. $\square$

Lemma 6 shows why an infinitely h-deep neighborhood can be $\|\cdot\|$-shallow. This property is well known from classical Gromov hyperbolic spaces, like the Klein distance on the open unit disc. But our set up is not Gromov hyperbolic, because $\partial\mathrm{P}$ might have faces. As a consequence there might be linear independent points in $\partial\mathrm{P}$ with finite h-distance. Nevertheless, our metric space $(\mathrm{P}^\circ \cap S, h)$ is “weakly” Gromov hyperbolic in the sense of [12]. The price we have to pay is that an infinitely h-deep neighborhood has to have a certain $\|\cdot\|$-width whenever P_1 is no single ray. Unfortunately, the shape of h-horoballs centered at points of $\partial\mathrm{P}$ is not known in our case. But Lemma 6 is sufficient for our purposes.

3.2. Spectral aspects

Like classical Perron–Frobenius theory our nonlinear version also has a spectral aspect. Instead of classical spectral notions we will use an interval calculus which is closer to (3) and is related to Collatz–Wielandt numbers [26].

An eigenvector of Λ is an $\mathcal{E} \in \mathrm{P} \setminus \{0\}$ such that $\Lambda(\mathcal{E}) = \lambda\mathcal{E}$ for some $\lambda \geqslant 0$. When there exists an eigenvector in $M \subset \mathrm{P}$ with eigenvalue λ, then we briefly say that M has an eigenvalue λ. The subset M is termed Λ-invariant when $\Lambda(M) \subset M$.

Lemma 7 (*Nussbaum [27, Lemma 3.1]*). *There exists an eigenvector in every closed, convex and* Λ*-invariant subcone of* D.

Proof (*Sketch*). Either an eigenvalue 0 exists or we can normalize Λ to stay on a suitable affine hyperplane. Now apply Brouwer’s fixed point theorem [34, Theorem 2.1.11]. $\square$

Lemma 8 (*Krein and Rutman [16, Theorem 9.1]*). *Suppose there exist* $\mathcal{A} \in \mathrm{D} \setminus \{0\}$ *and* $\lambda > 0$ *with* $\Lambda(\mathcal{A}) \geqslant \lambda\mathcal{A}$. *Then* D *has an eigenvalue not less than* λ.

Proof (*Sketch*). Since our map is superlinear, we can simplify the original proof. The set $\{\mathcal{A} \in \mathrm{D} | \Lambda(\mathcal{A}) \geqslant \lambda\mathcal{A}\}$ is a closed, convex and Λ-invariant cone by Proposition 2(ii) and (iii). Now apply Lemma 7. $\square$

Proposition 9. *Let* $\mathcal{A}, \mathcal{B} \in \mathrm{P} \setminus \{0\}$ *with* Λ*-invariant* $\mathrm{P}_{\mathcal{A}}$ *and* $\mathrm{P}_{\mathcal{B}}$.

(i) *When* $\ker \mathcal{A} \subset \ker \mathcal{B}$, *then* $\inf[\Lambda(\mathcal{B})/\mathcal{B}] \leqslant \sup[\Lambda(\mathcal{A})/\mathcal{A}]$.
(ii) *When* $\ker \mathcal{A} = \ker \mathcal{B}$, *then* $[\frac{\Lambda(\mathcal{A})}{\mathcal{A}}] \cap [\frac{\Lambda(\mathcal{B})}{\mathcal{B}}] \neq \emptyset$.
(iii) *The Collatz–Wielandt interval of* $\emptyset \neq M \subset \mathrm{P}_{\mathcal{A}}$, *given by*

$$\mathrm{cwi}(M) := \bigcap_{\mathcal{C} \in M} \left[\frac{\Lambda(\mathcal{C})}{\mathcal{C}} \right]$$

is a nonvoid compact subinterval of the positive reals.

Proof. (i) When $\ker \mathcal{A} \subset \ker \mathcal{B}$, then $\inf[\mathcal{A}/\mathcal{B}]$ is a positive real. Because of (4) and the assumed Λ-invariance there exists $g : V_0 \to \mathrm{R}$ with

$$\mathcal{A}(g) = \inf[\mathcal{A}/\mathcal{B}]\mathcal{B}(g) > 0,$$

$$\Lambda(\mathcal{A})(g) \geqslant \inf[\mathcal{A}/\mathcal{B}]\Lambda(\mathcal{B})(g) > 0.$$

(ii) Apply (i) twice.
(iii) $\mathrm{cwi}(M)$ is obviously compact and convex. Suppose it is void. Then

$$\sup_{\mathcal{C} \in M} \inf[\Lambda(\mathcal{C})/\mathcal{C}] > \inf_{\mathcal{C} \in M} \sup[\Lambda(\mathcal{C})/\mathcal{C}].$$

Each side can be approximated by a suitable sequence. Hence there exist $\mathcal{C}, \mathcal{C}' \in M$ such that $[\Lambda(\mathcal{C})/\mathcal{C}] \cap [\Lambda(\mathcal{C}')/\mathcal{C}'] = \emptyset$. This contradicts (ii). $\square$

Corollary 10. *Let* $n \in \mathrm{N} \setminus \{0\}$ *and* P_0 *be a* Λ*-invariant* P*-part. Then*

$$\mathrm{cwi}_n(\mathrm{P}_0) := \bigcap_{\mathcal{C} \in \mathrm{P}_0} \left[\frac{\Lambda^n(\mathcal{C})}{\mathcal{C}} \right]^{1/n}$$

is a nonvoid closed subinterval of $\mathrm{cwi}(\mathrm{P}_0) = \mathrm{cwi}_1(\mathrm{P}_0)$.

Proof. Let $\mathcal{A} \in \mathrm{P}_0$. Since P_0 is Λ-invariant, $[\Lambda^n(\mathcal{A})/\mathcal{A}]$ is defined for $\mathcal{A} \in \mathrm{P}_0$. For $f \notin \ker \mathrm{P}_0$ decompose $(\Lambda^n(\mathcal{A})/\mathcal{A})(f)$ into a product of quotients $(\Lambda^{k+1}(\mathcal{A})/\Lambda^k(\mathcal{A}))(f)$,

$0 \leqslant k < n$, and use Proposition 3 to see that

$$\inf[\Lambda(\mathcal{A})/\mathcal{A}]^n \leqslant \frac{\Lambda^n(\mathcal{A})}{\mathcal{A}}(f) \leqslant \sup[\Lambda(\mathcal{A})/\mathcal{A}]^n.$$

Thus $[\Lambda^n(\mathcal{A})/\mathcal{A}]^{1/n} \subset [\Lambda(\mathcal{A})/\mathcal{A}]$.

Since Λ^n also maps P to P, is positively homogeneous and superadditive, Proposition 9(iii) is also true for Λ^n instead of Λ. Thus for all $n \geqslant 1$, $\mathrm{cwi}_n(\mathrm{P}_0)$ is a compact nonvoid subinterval of $(0, \infty)$. $\square$

Corollary 10 is the reason why $\mathrm{cwi}(\mathrm{P}_0)$ is almost as good as an eigenvalue in P_0 even if the latter does not exist.

Corollary 11. *Let* $\mathrm{P}_1 \leqslant \mathrm{P}_0$ *be* Λ*-invariant* P*-parts and* $n \in \mathrm{N} \setminus \{0\}$. *Then*

$$\min \mathrm{cwi}_n(\mathrm{P}_1) \leqslant \max \mathrm{cwi}_n(\mathrm{P}_0).$$

Especially, when P_i *has an eigenvalue* λ_i *for* $i \in \{0, 1\}$, *then* $\lambda_1 \leqslant \lambda_0$ *and in the case of* $\mathrm{P}_0 = \mathrm{P}^\circ$ *every eigenvalue is not bigger than* λ_0.

Proof. This is a modification of [27, Lemma 3.3]. One applies Proposition 9(i) to all $\mathcal{A} \in \mathrm{P}_0$ and all $\mathcal{B} \in \mathrm{P}_1$. The result can be rewritten as $\min \mathrm{cwi}(\mathrm{P}_1) \leqslant \max \mathrm{cwi}(\mathrm{P}_0)$ according to Proposition 9(iii). Like in the proof of Corollary 10 we use that Λ^n and Λ have the same properties. This gives the cwi_n statement. $\square$

Corollary 11 can be used to estimate eigenvalues. In case of a strict inequality between the eigenvalues it tells us that P_0 must be strictly bigger than P_1. Estimates from below are facilitated by the next lemma.

Lemma 12. *Let* P_0 *be a* Λ*-invariant* P*-part. For every* $f \notin \ker \mathrm{P}_0$ *the map* $\mathcal{A} \mapsto \frac{\Lambda(\mathcal{A})}{\mathcal{A}}(f)$ *is quasi concave, that is, its upper level sets are convex. Especially, the map* $\mathcal{A} \mapsto \inf[\Lambda(\mathcal{A})/\mathcal{A}]$ *is quasi concave.*

Proof. The map $\mathcal{A} \mapsto (\Lambda(\mathcal{A}) - \lambda\mathcal{A})(f)$ is superadditive and positively homogeneous by Proposition 2(ii). This proves the first convexity statement. It implies the second convexity statement because the intersection of convex sets is convex. $\square$

Further spectral aspects of the cwi-calculus are given in Corollary 13, Proposition 33 and [18, Section 2].

Corollary 13. *When* P_0 *is a* P*-part with eigenvalue* λ, *then* $\mathrm{cwi}(\mathrm{P}_0) = \{\lambda\}$. *Especially, eigenvalues are unique on* P*-parts. When* P° *is* Λ*-invariant but* $\mathrm{P}^\circ \cap \mathrm{D}$ *has no eigenvalue, then* $\partial\mathrm{P} \cap \mathrm{D}$ *has an eigenvalue* $\max \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$.

Proof. (a) When P_0 has an eigenvalue λ, then Proposition 9(iii) implies $\mathrm{cwi}(M) = \{\lambda\}$ and the uniqueness of eigenvalues on P-parts.

(b) The remaining statement is [27, Theorem 3.1(3)]. We only sketch the arguments. When P° is Λ invariant, then $\mathrm{cwi}(\mathrm{P}^\circ)$ is defined. Corollary 11 shows that any eigenvalue of Λ is not bigger than $\max \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$. Let $\mathcal{M} \in \mathrm{D}^\circ$ be the Dirichlet form with conductances 1 on all edges of Γ_0 and consider the maps

$$\Lambda_m(\mathcal{A}) := \Lambda(\mathcal{A}) + \frac{\mathcal{M}}{m} \cdot \sum_{x \in V_0} \mathcal{A}(1_x)$$

for $m \in \mathrm{N} \setminus \{0\}$. They are positively homogeneous and superadditive, map $\mathrm{D} \setminus \{0\}$ into D° and $\mathrm{P} \setminus \{0\}$ into P°. Thus every such map has an eigenvector $\mathcal{E}^m \in \mathrm{D}^\circ$ with eigenvalue $\lambda_m > 0$ by Lemma 7. The argument in (a) shows that $(\lambda_m)_m$ decreases to a limit $\lambda \geqslant 0$. The local compactness of $(\mathrm{B}, \|\cdot\|)$ implies the existence of an accumulation point $\mathcal{E} \in \mathrm{D} \setminus \{0\}$. The sequence $(\Lambda_m)_m$ decreases uniformly to Λ on the unit $\|\cdot\|$-ball for $m \to \infty$. So $\Lambda(\mathcal{E}) = \lambda\mathcal{E}$. Now $\lambda \geqslant \max \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ follows from

$$\lambda_m = \max \mathrm{cwi}(\Lambda_m, \mathrm{P}^\circ \cap \mathrm{D}) \geqslant \max \mathrm{cwi}(\Lambda, \mathrm{P}^\circ \cap \mathrm{D}) \quad (m \geqslant 1). \qquad \square$$

Corollary 13 allows easy upper and lower estimates on eigenvalues. The uniqueness of eigenvalues was already obtained in [10, Corollary 3.7]. When $\mathrm{cwi}(\mathrm{P}_0) = \{\gamma\}$ then the corresponding eigenvector could unfortunately be located in the relative boundary of P_0.

3.3. *Dynamical aspects*

Like in classical Perron–Frobenius theory one can find eigenvectors of Λ by iteration of Λ on P [29]. The action of Λ on D-parts is new. Aspects of it were used implicitly in [32].

On a Λ-invariant P-part P_0 the normalized version of Λ,

$$\tilde{\Lambda}(\mathcal{A}) := \frac{\Lambda(\mathcal{A})}{\|\Lambda(\mathcal{A})\|}$$

is defined. For $\mathcal{A} \in \mathrm{P}_0$ the ω-limit set $\omega(\mathcal{A})$, the set of all accumulation points of $(\tilde{\Lambda}^n(\mathcal{A}))_n$, is $\|\cdot\|$-closed by definition and $\tilde{\Lambda}$-invariant when $\mathcal{A} \in \mathrm{D}$, because Λ is continuous on D.

Lemma 14 (*Nussbaum [28, Theorem 4.1]*). *Suppose the* P*-part* P_0 *contains an* $\mathcal{A}$ *such that* $(\Lambda^n(\mathcal{A}))_n$ *is h-bounded. Then* P_0 *contains an eigenvector.*

Proof (*Sketch*). The orbit is contained in an h-ball with radius r. Define $C := \bigcap_{\mathcal{C}\in\omega(\mathcal{A})} B^h_{2r}(\mathcal{C})$. This set is compact, convex and Λ-invariant. Apply Lemma 7. □

We know that D is Λ-invariant and closed. So we can start the orbit of Lemma 14 in $\mathrm{D}\cap\mathrm{P}^\circ$ and find an eigenvector in the same set. A stronger version of Lemma 14 is available, [28, Theorem 4.2], but we do not need it here.

Proposition 15 (*Metz* [25, Lemma 5]). *Let* P_0 *be a* Λ*-invariant* P*-part containing an* $\mathcal{A}\in\mathrm{D}$ *such that* $(\Lambda^n(\mathcal{A}))_n$ *is h-unbounded. Then* $(\tilde{\Lambda}_n(\mathcal{A}))_n$ *accumulates at a point* $\mathcal{B}\in\mathrm{D}$ *with* $\mathrm{P}_\mathcal{B} < \mathrm{P}_\mathcal{A}$ *and* $\mathrm{P}_\mathcal{B}\cap\mathrm{D}$ *contains an eigenvector.*

Proof (*Sketch*). Started in P_0 but h-unbounded, the normalized orbit must accumulate at $\overline{\mathrm{P}_0}\setminus\mathrm{P}_0$. The P-part of the accumulation point is Λ-invariant according to Lemma 5 and it is strictly smaller than P_0. Start a new normalized orbit in this P-part. It is either h-bounded or not. In the latter case iterate the argument. Every iteration increases the dimension of the kernel of the accumulation point strictly. But the accumulation point of a normalized orbit cannot be zero. So the new orbit must eventually be h-bounded. Now apply Lemma 14. □

The problem of detecting the P-parts $\mathrm{P}_\mathcal{B}$ of Proposition 15 is of the same type as the original eigenvalue problem (2) but with less dimensions. So we can use our methods recursively on simpler models.

Corollary 4 shows that Λ acts on the set of all P-parts supplemented by $\{0\}$. This mimics the analysis of closed classes of Markov chains, that is, we investigate the "positivity" of Λ, and we identify regions in which eigenvectors might be located. Another even finer aggregated action is the following. The property of coinciding graphs defines an equivalence relation on $\mathrm{D}\setminus\{0\}$. Let us call its equivalence classes D-parts. For $\mathcal{A},\mathcal{B}\in\mathrm{D}\setminus\{0\}$ denote the D-part of $\mathcal{A}$ by $\mathrm{D}_\mathcal{A}$. We will use $\mathrm{D}_\mathcal{B}\leqslant\mathrm{D}_\mathcal{A}$ to signify $\Gamma(\mathcal{B})\subset\Gamma(\mathcal{A})$, and $\mathrm{D}_\mathcal{B}<\mathrm{D}_\mathcal{A}$ when $\Gamma(\mathcal{B})$ is a strict subgraph of $\Gamma(\mathcal{A})$.

Lemma 16 (*Sabot* [32, Proposition 1.15]). *For* $\mathcal{A}\in\mathrm{D}$ *and different vertices* $x,y\in V_0$ *the conductance* $c_{\Lambda(\mathcal{A})}(x,y)$ *is positive if and only if there is a path in* $\Gamma(\mathcal{A}_1)$ *connecting* x *to* y *and avoiding* $V_0\setminus\{x,y\}$.

Proof (*Sketch*). Use the strong minimum principle, [13, Theorem 3.2.14], on connected components or employ Markov chains to prove the result. □

Lemma 16 shows that $\Gamma(\mathcal{A})\subset\Gamma(\mathcal{B})$ implies $\Gamma(\Lambda(\mathcal{A}))\subset\Gamma(\Lambda(\mathcal{B}))$. Hence Λ acts on the set of D-parts supplemented by $\{0\}$ and respects their partial ordering. The D-part action is finer than the P-part action, because a fixed collection of connected components can be realized by various graphs. The action can easily be calculated graphically via Lemma 16. Therefore, we will analyze the "positivity" of Λ via its D-part action.

On the Hany fractal every Dirichlet form has a unique collection of conductances $d_1,\ldots,d_4\in\mathrm{R}_+$. We index the D-parts by a representative conductance $(d_1,\ldots,d_4)\in$

$\{0,1\}^4$. In total there are $2^4-1=15$ D-parts, since $\{0\}$ is no D-part by definition. Obviously, $\mathrm{D}^\circ = \mathrm{D}_{(1,1,1,1)}$. The action of Λ on D-parts is given by

$$\begin{aligned} \Lambda : \{\mathrm{D}_{(1,0,0,0)}, \mathrm{D}_{(0,1,0,0)}, \mathrm{D}_{(1,1,0,0)}, \mathrm{D}_{(1,0,1,1)}\} &\mapsto \{\mathrm{D}_{(1,0,0,0)}\}, \\ \{\mathrm{D}_{(0,1,0,1)}, \mathrm{D}_{(1,1,0,1)}\} &\mapsto \{\mathrm{D}_{(1,1,0,1)}\}, \\ \{\mathrm{D}_{(0,1,1,0)}, \mathrm{D}_{(1,1,1,1)}\} &\mapsto \{\mathrm{D}_{(1,1,1,1)}\}, \\ \{\mathrm{D}_{(0,0,1,0)}, \mathrm{D}_{(0,0,0,1)}, \mathrm{D}_{(0,0,1,1)}\} &\mapsto \{0\}. \end{aligned}$$

Since Λ respects the partial ordering of D-parts, the remaining cases can be deduced from the above list. To summarize, only the D-parts D°, $\mathrm{D}_1 := \mathrm{D}_{(1,0,0,0)}$ and $\mathrm{D}_3 := \mathrm{D}_{(1,1,0,1)}$ are Λ-invariant. We would like to find an eigenvector in D° but we want to avoid accumulation points in D_1 or D_3. The connected components of the reducible D-parts are

$$\begin{aligned} &\mathrm{D}_1 : \{v_2\}, \{v_1, v_3\}, \{v_4\}, \{v_5\}, \{v_6\}, \{v_7\}, \\ &\mathrm{D}_3 : \{v_1, v_2, v_3\}, \{v_4, v_6\}, \{v_5, v_7\}. \end{aligned} \tag{5}$$

Let us denote the P-part containing D_i by P_i, for $i=1,3$. We see that $\mathrm{P}_1 < \mathrm{P}_3$. In the above discussion of the Λ-action on D-parts we have cut certain edges off in choosing their conductances to be 0 instead of positive. This is the origin of the term "cut" in the title of this article.

Formally, the Hany fractal is not connected and thus not contained in the class of fractals considered in [32]. But one might try to apply Sabot's methods to the two connected components seperately. Even then his Theorem 5.1 cannot be employed, because his Assumption H is violated by $\mathrm{P}_1 < \mathrm{P}_3$! This is the main reason why we have to improve Sabot's theorem.

To see whether D_1 or D_3 contain accumulation points, we have to find eigenvectors inside them according to Proposition 15. Obviously, D_1 consists of eigenvectors with eigenvalue $\frac{1}{3}$. Furthermore, the function $h := 1_{v_1} - 1_{v_3}$ is an element of $\operatorname{im}\mathrm{D}_1$ satisfying

$$\Lambda(\mathcal{A})(h) = \tfrac{1}{3}\,\mathcal{A}(h) \text{for all } \mathcal{A} \in \mathrm{D} \text{ with } \mathrm{D}_1 \leqslant \mathrm{D}_{\mathcal{A}} \tag{6}$$

[24, Section 4]. So the eigenvalues of D_3 and D° equal $\frac{1}{3}$ provided they exist at all. The eigenvalue test in [25, Theorem 1] assumes that the eigenvalue of D° is strictly bigger than the one of D_3. So this test does not apply either! For the same reason the Gâteaux derivative test in [29, Theorem 3.3] does not apply! Because of all these negative remarks the Hany fractal is a good bench mark test for existence criteria.

4. Monotone convergence

To calculate directional derivatives of Λ at $\partial \mathrm{P} \cap \mathrm{D}$ in Section 5 we have to consider sequences $(\Lambda(\mathcal{D}_k + n\mathcal{B}_k))_n$ for

$$k \in \mathrm{N}, \mathcal{B} \in \mathrm{D} \setminus \{0\} \text{ and } \mathcal{D} \in \mathrm{B} \text{ such that } \mathcal{B} + \mathcal{D} \in \mathrm{D}. \tag{7}$$

Since Λ is P-increasing the limit can be defined via monotone convergence. Electrically, the infinite conductances of $\infty\mathcal{B}$ short circuit $\Gamma(\mathcal{B}_k)$. This is the origin of the term "short" in the title of this article.

For $f : V_k \to \mathrm{R}$ the sequence $((\mathcal{D}_k + n\mathcal{B}_k)(f))_n$ is increasing and the limit is finite if and only if $f \in \ker \mathcal{B}_k$. This means f is constant on the connected components of $\Gamma(\mathcal{B}_k)$, also called $\mathcal{B}_k$-components for short.

Lemma 17. *Let $k \in \mathrm{N}$ and $P := \{C_1, \ldots, C_l\}$ be a partition of V_k. Choose $g : V_k \to \mathrm{R}$ to be constant on each C_i, $1 \leqslant i \leqslant l$. Define $\overline{g} : P \to \mathrm{R}$ to be $\overline{g}(C_i) := g(x)$ for some $x \in C_i$, $1 \leqslant i \leqslant l$. For $\mathcal{A} \in \mathrm{P}$ let $\overline{c}_{\mathcal{A}_k} : P \times P \to \mathrm{R}$ be given by*

$$\overline{c}_{\mathcal{A}_k}(C_i, C_j) := \begin{cases} 0, & i = j, \\ \sum_{x \in C_i, y \in C_j} c_{\mathcal{A}_k}(x, y), & i \neq j. \end{cases}$$

Then

$$\mathcal{A}_k(g) = \frac{1}{2} \sum_{i,j=1}^{l} (\overline{g}(C_j) - \overline{g}(C_i))^2 \overline{c}_{\mathcal{A}_k}(C_i, C_j). \tag{8}$$

Proof. Wright $\mathcal{A}_k(g)$ as a quadratic form with coefficients $c_{\mathcal{A}_k}(x, y)$ for $x, y \in V_k$. Now use the properties of a partition. $\square$

In the situation of (7) let us apply Lemma 17 to $\mathcal{A}_k := \mathcal{D}_k + n\mathcal{B}_k$ and $g \in \ker \mathcal{B}_k$. Then g is constant on $\mathcal{B}_k$-components which we define to be the elements of P. Denote the Dirichlet form on P defined via the conductance $\overline{c}_{\mathcal{A}_k}$ by $\overline{\mathcal{A}}_k$. Then (8) implies

$$(\mathcal{D}_k + \infty\mathcal{B}_k)(g) = \overline{\mathcal{A}}_k(\overline{g}). \tag{9}$$

We define the domain of $\mathcal{D}_k + \infty\mathcal{B}_k$ to be $\ker \mathcal{B}_k$, because it is finite on exactly this set. The $\overline{(\cdot)}$-map in Lemma 17 maps $\ker \mathcal{B}_k$ to R^P the domain of $\overline{\mathcal{A}}_k$. We have reduced the analysis of Dirichlet forms with infinite conductances to the study of ordinary Dirichlet forms.

Physically, we have short circuited all edges of $\Gamma(\mathcal{B}_k)$ inside the electrical resistor network $(V_k, \mathcal{B}_k)$. Mathematically, we use a calculus for infinite conductances based on (9). This means we have to use the standard conventions of measure theory regarding the ∞-calculus. For $\mathcal{C} \in \mathrm{D}_{\mathcal{B}}$ we have $\mathcal{D}_k + \infty\mathcal{B}_k = \mathcal{D}_k + \infty\mathcal{C}_k$. Therefore we define $\mathcal{D}_k +$

$\infty\mathcal{B}_k =: \mathcal{D}_k + \infty \mathrm{D}_{\mathcal{B}_k}$. Since $\Gamma(\mathcal{D}_k + \infty \mathrm{D}_{\mathcal{B}_k})$ does not depend on $\mathcal{D}$ as long as $\mathcal{D} \in \mathrm{D} \setminus \{0\}$ stays in a fixed D-part, we also allow the notation $\Gamma(\Psi^k(\mathrm{D}_{\mathcal{D}}) + \infty \Psi^k(\mathrm{D}_{\mathcal{B}})) := \Gamma(\mathcal{D}_k + \infty\mathcal{B}_k)$. The graph $\Gamma(\mathrm{D}^\circ + \infty \mathrm{D}_1)$ of our Hany fractal is shown in the third column of Fig. 2. Its graph $\Gamma(\mathrm{D}^\circ + \infty \mathrm{D}_3)$ is depicted in the fourth column of the same figure.

For $\mathcal{A} \in \mathrm{P}$ and $M \subset V_0$ define the restricted Dirichlet form $\mathcal{A}^M$ for $f : M \to \mathrm{R}$ by

$$\mathcal{A}^M(f) := \frac{1}{2} \sum_{x,y \in M} (f(y) - f(x))^2 c_{\mathcal{A}}(x, y).$$

For $\mathcal{A} \in \mathrm{D}$ and $k \in \mathrm{N}$ let us call an $\mathcal{A}_k$-component intersecting V_0 a boundary $\mathcal{A}_k$-component and an interior $\mathcal{A}_k$-component otherwise. In the situation of (7) we denote the union of all boundary $\mathcal{B}_1$-components by $C_0^{\mathcal{B}}$. For $f : V_0 \to \mathrm{R}$ consider

$$\inf\{\mathcal{D}_1(g) | g \in \mathrm{R}^{V_1}, g|_{C_0^{\mathcal{B}}} = H^{\mathcal{B}_1}_{V_1 \setminus V_0} f, g|_{C_i} \text{ constant for } 1 \leqslant i \leqslant l\}. \tag{10}$$

By Lemma 17 the minimizing element is unique when there are no interior $(\mathcal{B} + \mathcal{D})_1$-components. We define it to have a minimal $L^2(V_1)$-norm otherwise. Let us denote this specific minimizer of (10) by $H^{(\infty\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f$ because of the following proposition.

Proposition 18. *Let $\mathcal{B} \in \mathrm{D} \setminus \{0\}$ and $\mathcal{D} \in \mathrm{B}$ such that $\mathcal{B} + \mathcal{D} \in \mathrm{D}$. For $f : V_0 \to \mathrm{R}$:*

$$\lim_{n\to\infty} H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f = H^{(\infty\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f.$$

Proof. Because of $\mathcal{D} + n\mathcal{B} \in \mathrm{D}$ and the minimum principle, $|H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f(x)|$ is bounded above by $\|f\|_{L^\infty(V_0)}$. Hence there exists an accumulation point $a \in \mathrm{R}^{V_1}$ of $(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f)_n$. For every $n \in \mathrm{N} \setminus \{0\}$ the Dirichlet principle shows

$$H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f = H^{(\mathcal{B}+\mathcal{D}/n)_1}_{V_1 \setminus V_0} f.$$

Thus a is also an accumulation point of $(H^{(\mathcal{B}+\mathcal{D}/n)_1}_{V_1 \setminus V_0} f)_n$. The convergence of $(\mathcal{B}+\mathcal{D}/n)_1$ to $\mathcal{B}_1$ is locally uniform in $L^\infty(V_1)$. Hence a minimizes $\mathcal{B}_1(g)$ among all $g : V_1 \to \mathrm{R}$ with $g|_{V_0} = f$. That is, $a \in H^{\mathcal{B}_1}_{V_1 \setminus V_0} f + \ker \mathcal{B}_1$. Set $H^{\mathcal{B}_1}_{V_1 \setminus V_0} f =: h$. According to the minimum principle $a = h$ on $C_0^{\mathcal{B}}$ and h is constant on all interior $\mathcal{B}_1$-components. Again by the minimum principle

$$\|H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f - H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus C_0^{\mathcal{B}}} h\|_{L^\infty(V_1)} \leqslant \|H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f - h\|_{L^\infty(C_0^{\mathcal{B}})}.$$

Thus a is also an accumulation point of $(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus C_0^{\mathcal{B}}} h)_n$.

We abbreviate $H^{(\infty\mathcal{B}+\mathcal{D})_1}_{V_1\setminus V_0} f$ by m. The Dirichlet principle shows

$$(n\mathcal{B}+\mathcal{D})_1(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1\setminus C_0^{\mathcal{B}}} h) \leqslant (n\mathcal{B}+\mathcal{D})_1(m). \tag{11}$$

Since $\Gamma(\mathcal{B}_1)$ does not connect $C_0^{\mathcal{B}}$ to $V_1 \setminus C_0^{\mathcal{B}}$, we have $\mathcal{B}_1 = \mathcal{B}_1^{C_0^{\mathcal{B}}} + \mathcal{B}_1^{V_1\setminus C_0^{\mathcal{B}}}$. Noticing that $h = m$ on $C_0^{\mathcal{B}}$ and that m is constant on interior $\mathcal{B}_1$-components, we deduce

$$\mathcal{B}_1^{C_0^{\mathcal{B}}}(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1\setminus C_0^{\mathcal{B}}} h) = \mathcal{B}_1^{C_0^{\mathcal{B}}}(m) \text{ and } \mathcal{B}_1^{V_1\setminus C_0^{\mathcal{B}}}(m) = 0.$$

Incorporation this into (11) we conclude

$$\mathcal{D}_1(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1\setminus C_0^{\mathcal{B}}} h) \leqslant \left(n\mathcal{B}_1^{V_1\setminus C_0^{\mathcal{B}}} + \mathcal{D}_1\right)(H^{(n\mathcal{B}+\mathcal{D})_1}_{V_1\setminus C_0^{\mathcal{B}}} h) \leqslant \mathcal{D}_1(m).$$

This implies $\mathcal{D}_1(a) \leqslant \mathcal{D}_1(m)$. On the other hand, a is a candidate function in the variational problem (10), that is, $\mathcal{D}_1(m) \leqslant \mathcal{D}_1(a)$. Thus $m - a \in \ker \mathcal{D}_1$. By the minimum principle $m = a$ on boundary $(\mathcal{B}+\mathcal{D})_1$-components. On interior $(\mathcal{B}+\mathcal{D})_1$-components both functions are defined to be 0. Hence $m = a$. $\square$

Now it is easy to calculate $\Lambda(\infty\mathcal{B} + \cdot)$.

Corollary 19. *Let $\mathcal{B} \in \mathrm{D}\setminus\{0\}$ and $\mathcal{D} \in \mathrm{B}$ with $\mathcal{B}+\mathcal{D} \in \mathrm{D}$. Then $\Lambda(\mathcal{D}+\infty\mathcal{B})(f)$ is finite if and only if $f \in \ker \Lambda(\mathcal{B})$. For such an f,*

$$\Lambda(\mathcal{D}+\infty\mathcal{B})(f) = (\mathcal{D}+\infty\mathcal{B})_1(H^{(\mathcal{D}+\infty\mathcal{B})_1}_{V_1\setminus V_0} f).$$

When P° and $\mathrm{P}_{\mathcal{B}}$ are Λ-invariant, then $\ker(\infty\mathcal{B} + \mathrm{P}^\circ) = \ker \mathrm{P}^\circ$.

Proof. Proposition 18 shows that the asserted equality is true even when the participating energies are infinite. We know that $(\infty\mathcal{B}+\mathcal{D})_1(g) < \infty$ if and only if $g \in \ker \mathcal{B}_1$. But $H^{(\mathcal{D}+\infty\mathcal{B})_1}_{V_1\setminus V_0} f \in \ker \mathcal{B}_1$ if and only if it is constant on every boundary $\mathcal{B}_1$-component. This happens if and only if $f \in \ker \Lambda(\mathcal{B})$ according to Lemma 16.

Let $\mathcal{A} \in \mathrm{P}^\circ$. Then $\Lambda(\infty\mathcal{B} + \mathcal{A})$ is finite on $\ker \Lambda(\mathcal{B}) = \ker \mathcal{B}$. Note that $\ker \Lambda(\mathcal{A})$ equals $\ker \mathcal{A}$ which coincides with $\ker \mathrm{P}^\circ$. For $f \notin \ker \mathrm{P}^\circ$ we have

$$\Lambda(\infty\mathcal{B}+\mathcal{A})(f) \geqslant \Lambda(\mathcal{A})(f) > 0.$$

Thus $\ker \Lambda(\infty\mathcal{B}+\mathcal{A}) \subset \ker \mathrm{P}^\circ$. Suppose $0 = \Lambda(\mathcal{A})(f) < \Lambda(\infty\mathcal{B}+\mathcal{A})(f)$. Then $H^{\mathcal{A}_1}_{V_1\setminus V_0} f \in \ker \mathcal{A}_1$ and, because of $\ker \mathcal{A} \subset \ker \mathcal{B}$, we have $\ker \mathcal{A}_1 \subset \ker \mathcal{B}_1$. Thus

$$0 = \mathcal{A}_1(H^{\mathcal{A}_1}_{V_1\setminus V_0} f) = (\infty\mathcal{B}+\mathcal{A})_1(H^{\mathcal{A}_1}_{V_1\setminus V_0} f) < (\infty\mathcal{B}+\mathcal{A})_1(H^{(\infty\mathcal{B}+\mathcal{A})_1}_{V_1\setminus V_0} f).$$

This contradicts the Dirichlet principle. $\square$

Let $0 \neq \mathcal{B} \in \partial \mathrm{P} \cap \mathrm{D}$. In $\mathcal{D} + \infty\mathcal{B}$ all edges of $\Gamma(\mathcal{B})$ have infinite conductances thus only those conductances of $\mathcal{D}$ matter which do not connect points of a $\mathcal{B}$-component. Denote these components by $C_1, \ldots, C_l$ and define for $\mathcal{A} \in \mathrm{P}$, $f : V_0 \to \mathrm{R}$ the projection

$$\Pi_1(\mathcal{A})(f) := \sum_{i=1}^{l} \mathcal{A}^{C_i}(f)$$

and for $f \in \ker \mathcal{B}$ the projection

$$\Pi_\infty(\mathcal{A})(f) := (\mathcal{A} + \infty\mathcal{B})(f) = \mathcal{A}(f). \tag{12}$$

To avoid confusion we occasionally use $\Pi_{\mathcal{B}}$ or $\Pi_{\infty\mathcal{B}}$ instead of Π_1 or Π_∞ to stress the dependence on $\mathcal{B}$. Denote the set of all $\mathcal{B}$-components by P, the cone of all Dirichlet forms on P by $\mathrm{D}_\infty = \mathrm{D}_{\infty\mathcal{B}}$ and its cone of positive semidefinite forms in $\mathrm{D}_\infty - \mathrm{D}_\infty$ by $\mathrm{P}_\infty = \mathrm{P}_{\infty\mathcal{B}}$.

Lemma 20. *Let* $0 \neq \mathcal{B} \in \partial \mathrm{P} \cap \mathrm{D}$. *When* P° *is* Λ*-invariant then* P°_∞ *also is, and for all* $\mathcal{D} \in \mathrm{B}$ *with* $\mathcal{B} + \mathcal{D} \in \mathrm{P}^\circ$ *and all* $f \in \ker \mathcal{B} \setminus \ker \mathrm{P}^\circ$,

$$\frac{\Lambda^n(\infty\mathcal{B} + \mathcal{D})}{\infty\mathcal{B} + \mathcal{D}}(f) \geqslant \frac{\Lambda^n(\mathcal{B} + \mathcal{D})}{\mathcal{B} + \mathcal{D}}(f).$$

Proof. The choice of f implies $\infty\mathcal{B}(f) = \mathcal{B}(f) = 0$. Now the inequality follows by monotonicity. $\square$

5. Gâteaux derivatives

The discussion after Corollary 4 explains why we have to rely on superadditivity when we want Λ to contract h-distances. How can we detect strict superadditivity efficiently? We will explain why Gâteaux derivatives of Λ at $\partial \mathrm{P} \cap \mathrm{D}$ are powerful tools.

For $\mathcal{A}, \mathcal{B} \in \mathrm{P}$, $f : V_0 \to \mathrm{R}$ and $\alpha \geqslant 0$ the map $\alpha \mapsto \Lambda(\mathcal{A} + \alpha\mathcal{B})(f)$ is concave in the classical sense. Thus its additivity on $\mathcal{A} + \mathcal{B}$ implies the additivity on the cone spanned by $\mathcal{A}$ and $\mathcal{B}$ [35, Theorem 5.1.8]. But even more is true.

Lemma 21 (*Metz [23, Corollary 4(c)]*). *Let* $\mathcal{A}, \mathcal{B} \in \mathrm{P}$, $f : V_0 \to \mathrm{R}$ *and suppose* $\Lambda(\mathcal{A} + \mathcal{B})(f)$ *equals* $\Lambda(\mathcal{A})(f) + \Lambda(\mathcal{B})(f)$. *Then* $\Lambda(\cdot)(f)$ *is linear on the intersection of* P *with the vector space spanned by* $\mathcal{A}$ and $\mathcal{B}$.

Proof (*Sketch*). By Proposition 2(ii) the additivity condition implies that $H^{\mathcal{A}_1}_{V_1 \setminus V_0} f$ is $\mathcal{B}_1$-harmonic on $V_1 \setminus V_0$. This turns the trace $\mathrm{Tr}(\cdot)(f)$ into a linear function on the intersection of P with the space spanned by $\mathcal{A}$ and $\mathcal{B}$. The coupling map is linear anyway. $\square$

According to Lemma 21 we can detect linearity anywhere on a line segment in P. Since the one sided derivatives of a concave function are monotone, a linearity test based on derivatives is easiest when it is carried out at the extremal points of the line segment. This suggests to use derivatives at ∂P provided they exist.

Lemma 22. *Let* $0 \neq \mathcal{B} \in \partial \mathrm{P} \cap \mathrm{D}$, $f : V_0 \to \mathrm{R}$ *and* $\mathcal{D} \in \mathrm{B}$ *such that* $\mathcal{B} + \mathcal{D} \in \mathrm{D}$. *Then the following statements hold*:

(i) *The right sided derivative of* $\varepsilon \mapsto \Lambda(\mathcal{B} + \varepsilon\mathcal{D})(f)$ *exists in* $\varepsilon = 0$. *It is*

$$\frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}(f) := \lim_{\varepsilon \downarrow 0} D\Lambda_{\mathcal{B}+\varepsilon\mathcal{D}}(\mathcal{D})(f) = \mathcal{D}_1(H^{(\infty\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f).$$

The derivative is continuous in $\mathcal{D} \in \mathrm{B}$ *as long as* $\mathcal{B} + \mathcal{D} \in \mathrm{D}$.

(ii) $\frac{\Lambda(\mathcal{B}+\varepsilon\mathcal{D}) - \Lambda(\mathcal{B})}{\varepsilon}(f)$ *is decreasing in* $\varepsilon \in (0, 1)$.

(iii) $\Lambda(\mathcal{B} + \varepsilon\mathcal{D}) \leqslant \Lambda(\mathcal{B}) + \varepsilon \cdot \frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}$ *for* $\varepsilon \in [0, 1)$.

(iv) *For* $\varepsilon \downarrow 0$ *the quotient in* (ii) *converges to the derivative in* (i) *uniformly in* $\mathcal{D}$ *on* $\|\cdot\|$*-bounded sets and simultaneously uniformly in* $\mathcal{B}$ *on* $\|\cdot\|$*-compact subsets of a fixed* P*-part.*

Proof. (i) By Proposition 2 the map $\varepsilon \mapsto \Lambda(\mathcal{B} + \varepsilon\mathcal{D})(f)$ is continuous in $[0, 1]$, continuously differentiable in $(0, 1)$ and concave in $(0, 1)$. Thus its one sided derivatives in $(0, 1)$ are monotone [35, Corollary 5.1.4]. The mean value theorem of differential calculus now implies the existence of $\frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}(f)$ as the monotone, possibly infinite, limit in (i). By Propositions 2(iii) and 18 the limit has the desired value.

Define a partition P of V_1 in requiring every point of a boundary $\mathcal{B}_1$-component to form its own one point set and every interior $\mathcal{B}_1$-component to form a seperate set. Denote the collection of sets arising from boundary $\mathcal{B}_1$-components by P_0. Then Lemma 17 allows to rewrite (10) as

$$\frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}(f) = \mathrm{Tr}_{P_0}(\mathcal{D}_1)(H^{\mathcal{B}_1}_{V_1 \setminus V_0} f).$$

The proof of $\Lambda \in C(\mathrm{D})$ in Proposition 2(iii) shows that every trace is $\|\cdot\|$-continuous on Dirichlet forms. Hence $\mathcal{D} \mapsto \frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}$ is continuous in $\mathcal{D} \in \mathrm{D}$. To treat $\mathcal{D} \in \mathrm{B}$ such that $\mathcal{B} + \mathcal{D} \in \mathrm{D}$, we only use the formula of the derivative to realize that

$$\frac{\partial}{\partial(\mathcal{B} + \mathcal{D})} \Lambda_{\mathcal{B}}(f) = \Lambda(\mathcal{B})(f) + \frac{\partial}{\partial \mathcal{D}} \Lambda_{\mathcal{B}}(f).$$

(ii) This monotonicity is a classical property of convex functions [35, Corollary 5.1.2].

(iii) This statement is true in P° by Proposition 2. Using (i) and the continuity of Λ on D, it carries over to $\mathcal{B} \in \partial\mathrm{P} \cap \mathrm{D}$.

(iv) Let B be the unit ball in $L^2(V_0)$, K be a $\|\cdot\|$-compact subset of a fixed P-part intersected with D, and N a $\|\cdot\|$-compact subset of B. Then

$$C := \{(f, \mathcal{B}, \mathcal{D}) \in B \times K \times N | \mathcal{B} + \mathcal{D} \in \mathrm{D}\}$$

is compact in the product topology of $B \times K \times N$, because D is $\|\cdot\|$-closed. The map $(f, \mathcal{B}, \mathcal{D}) \mapsto \frac{\Lambda(\mathcal{B}+\varepsilon\mathcal{D})-\Lambda(\mathcal{B})}{\varepsilon}(f)$ is continuous on C since the $\|\cdot\|$-convergence of forms is uniform on B. The map $(f, \mathcal{B}, \mathcal{D}) \mapsto \frac{\partial}{\partial\mathcal{D}}\Lambda_{\mathcal{B}}(f)$ is continuous on C by (i). Now (ii) and Dini's theorem, [11, Satz 108.1] imply the desired uniform convergence. $\square$

For $f \in \ker\mathcal{B}$ the harmonic continuation $H_{V_1\setminus V_0}^{(\infty\mathcal{B}+\mathcal{D})_1} f$ is constant on all $\mathcal{B}_1$-components by (10). Hence

$$\begin{aligned}\frac{\partial}{\partial\mathcal{D}}(f) &= \mathcal{D}_1(H_{V_1\setminus V_0}^{(\infty\mathcal{B}+\mathcal{D})_1} f) \\ &= (\infty\mathcal{B} + \mathcal{D})_1(H_{V_1\setminus V_0}^{(\infty\mathcal{B}+\mathcal{D})_1} f) \\ &= \Lambda(\infty\mathcal{B} + \mathcal{D})(f).\end{aligned} \tag{13}$$

This is the shorting interpretation of the Gâteaux derivatives when they are evaluated at $\ker\mathcal{B}$. The nonlinearity in this statement shows that the derivative in Lemma 22(i) is Gâteaux but not Fréchet.

For $0 \neq \mathcal{B} \in \partial\mathrm{P} \cap \mathrm{D}$ we have $\mathrm{R}^{V_0} = \ker\mathcal{B} \oplus \operatorname{im}\mathcal{B}$. We decompose $u : V_0 \to \mathrm{R}$ accordingly into $f + g$. Suppose $\mathrm{P}_{\mathcal{B}}$ is Λ-invariant. Then the scaling of u near $\mathcal{B}$ in direction $\mathcal{D} \in \mathrm{P}$ is

$$\lim_{\varepsilon\downarrow 0} \frac{\Lambda(\mathcal{B} + \varepsilon\mathcal{D})}{\mathcal{B} + \varepsilon\mathcal{D}}(u) = \frac{\Lambda(\mathcal{B})}{\mathcal{B}}(g) \tag{14}$$

by the continuity of Λ on D. Thus the scaling is completely described by $[\Lambda(\mathcal{B})/\mathcal{B}]$. It remains to look at $f \in \ker\mathcal{B} \setminus \ker\mathrm{P}^\circ$. According to the rule of l'Hospital and (13),

$$\lim_{\varepsilon\downarrow 0} \frac{\Lambda(\mathcal{B} + \varepsilon\mathcal{D})}{\mathcal{B} + \varepsilon\mathcal{D}}(f) = \frac{\frac{\partial}{\partial\mathcal{D}}\Lambda_{\mathcal{B}}}{\mathcal{D}}(f) = \frac{\Lambda(\infty\mathcal{B} + \mathcal{D})}{\infty\mathcal{B} + \mathcal{D}}(f). \tag{15}$$

Thus the scaling is completely described by $[\Lambda(\infty\mathcal{B}+\mathcal{D})/(\infty\mathcal{B}+\mathcal{D})]$. For the above f Lemma 22(ii) turns into: $\varepsilon \mapsto \frac{\Lambda(\mathcal{B}+\varepsilon\mathcal{D})}{\varepsilon\mathcal{D}}(f)$ is decreasing in $\varepsilon \in (0, 1)$. Thus the scaling at $\varepsilon = 0$ is maximal! This is the technical reason why Gâteaux derivatives at $\partial\mathrm{P} \cap \mathrm{D}$ are very effective with respect to scaling arguments.

Eqs. (14) and (15) advise us to study $\Lambda \circ \Pi_1$ and $\Lambda \circ \Pi_\infty$.

Proposition 23 (*Sabot [32, Proposition 4.23]*). *Let* $0 \neq \mathcal{B} \in \partial \mathrm{P} \cap \mathrm{D}$ *and suppose that* P° *and* $\mathrm{P}_{\mathcal{B}}$ are Λ*-invariant. For every* $\varepsilon > 0$ *there exists a* $\|\cdot\|$*-neighborhood* U *of* $\mathrm{P}_{\mathcal{B}}$ *such that for all* $\mathcal{A} \in U \cap \mathrm{D} \cap \mathrm{P}^\circ$,

$$\left[\frac{\Pi_1 \circ \Lambda(\mathcal{A})}{\Lambda \circ \Pi_1(\mathcal{A})}\right] \subset [1-\varepsilon, 1+\varepsilon] \text{ and } \left[\frac{\Pi_\infty \circ \Lambda(\mathcal{A})}{\Lambda \circ \Pi_\infty(\mathcal{A})}\right] \subset [1-\varepsilon, 1]. \tag{16}$$

Proof. We give a proof based on $\Lambda \in C(\mathrm{D})$ and Gâteaux derivatives.

First inclusion in (16): For $x, y \in V_0$ the map $\mathcal{C} \mapsto \mathcal{C}(1_x, 1_y)$ is $\|\cdot\|$-continuous. Thus $U' := \{\mathcal{C} \in \mathrm{D} | \Gamma(\mathcal{B}) \subset \Gamma(\mathcal{C})\}$ is a $\|\cdot\|$-neighborhood of $\mathrm{P}_{\mathcal{B}}$ in D. Obviously $\Pi_1(U') \subset \mathrm{P}_{\mathcal{B}}$. Let $\mathcal{A} \in U'$. Since Λ respects the partial ordering of D-parts and $\mathrm{P}_{\mathcal{B}}$ is Λ-invariant, the numerator and the denominator are elements of $\mathrm{P}_{\mathcal{B}}$. Thus the range has to be calculated on $\operatorname{im} \mathcal{B}$. Since $\mathrm{P}_{\mathcal{B}}$ is Λ-invariant and Π_1 is a projection to $\mathrm{P}_{\mathcal{B}}$, the quotient in question is 1 on $\mathrm{P}_{\mathcal{B}} \cap \mathrm{D}$. We know $\Pi_1, \Lambda \in C(\mathrm{D})$ and the $\|\cdot\|$-limits are locally uniform on $\operatorname{im} \mathcal{B}$. Thus there exists a $\|\cdot\|$-neighborhood $U'' \subset U'$ on which the first inclusion holds.

Second inclusion in (16): Let $\mathcal{A} \in \mathrm{P}^\circ$ and $f \in \ker \mathcal{B} \setminus \ker \mathrm{P}^\circ$. Since P° is Λ-invariant, $\Lambda(\mathcal{A}) \in \mathrm{P}^\circ$ and

$$\Pi_\infty \circ \Lambda(\mathcal{A})(f) = (\infty \mathcal{B} + \Lambda(\mathcal{A}))(f) = \Lambda(\mathcal{A})(f) > 0. \tag{17}$$

So the numerator is defined on $\ker \mathcal{B}$ and its kernel is $\ker \mathrm{P}^\circ$. By Corollary 19 the denominator is defined on $\ker \mathcal{B}$ and has the kernel P°. Because of (13),

$$\Lambda \circ \Pi_\infty(\mathcal{A})(f) = \frac{\partial}{\partial(\mathcal{A} - \Pi_1(\mathcal{A}))} \Lambda_{\Pi_1(\mathcal{A})}(f). \tag{18}$$

Because of (17) and (18), Lemma 22(iii) takes the obvious form $\Lambda(\mathcal{A}) \leqslant \Lambda(\infty \mathcal{B} + \mathcal{A})$. Consequently

$$\left[\frac{\Pi_\infty \circ \Lambda(\mathcal{A})}{\Lambda \circ \Pi_\infty(\mathcal{A})}\right] \subset (0, 1] \quad (\mathcal{A} \in \mathrm{P}^\circ).$$

The remaining assertion is now a consequence of Lemma 22(iv). □

Proposition 23 says that a single commutation of Λ and Π_i causes a multiplicative factor, $i \in \{1, \infty\}$. Since Λ and Π_i are positively homogeneous, we can exchange repeatedly. For $n \in \mathrm{N} \setminus \{0\}$ and $1 \leqslant k \leqslant n$ this results in:

$$\left[\frac{\Pi_1 \circ \Lambda^k(\mathcal{A})}{\Lambda^k \circ \Pi_1(\mathcal{A})}\right] \subset [1-\varepsilon, 1+\varepsilon]^k \text{ and } \left[\frac{\Pi_\infty \circ \Lambda^k(\mathcal{A})}{\Lambda^k \circ \Pi_\infty(\mathcal{A})}\right] \subset [1-\varepsilon, 1]^k, \tag{19}$$

where $\mathcal{B}$, $\mathrm{P}_{\mathcal{B}}$, ε, and U are as in Proposition 23 and where we assume that $\{\Lambda^k(\mathcal{A}) | 0 \leqslant k \leqslant n\} \subset U \cap \mathrm{D} \cap \mathrm{P}^\circ$. Alternatively one could use the chain rule to prove (19).

6. Existence

The surprising fundamental observation of this section is that the dynamics of Λ near $\partial \mathrm{P} \cap \mathrm{D}$ has a one dimensional flavor because of the cwi-calculus and the shorting interpretation of Gâteaux derivatives. We will try to follow a well known argument: The fixed point 0 of a continuous C^1-map on the unit interval is repellent, when the one sided derivative at zero is strictly bigger than one [31, p. 22]. Nussbaum has used Gâteaux derivatives in [29, Section 3]. But we have seen already that his results cannot be applied to the Hany fractal. Therefore, we have to refine his methods. One possibility is to follow Sabot's idea in [32] which is roughly speaking: For $\mathcal{A} \in \mathrm{P}^\circ$ near an eigenvector $\mathcal{B} \in \partial \mathrm{P}$ the scaling $(\Lambda(\mathcal{A})/\mathcal{A})(f)$ of a function $f \notin \ker \mathcal{B}$ is almost the eigenvalue of $\mathcal{B}$. When the scaling of a function $g \in \ker \mathcal{B}$ stays uniformly above the eigenvalue of $\mathcal{B}$, then a $\tilde{\Lambda}$-orbit cannot accumulate at $\mathcal{B}$ because in the limit $\mathcal{B}(g) = 0 < \mathcal{B}(f)$. The rule of l'Hospital for Gâteaux derivatives in (15) will make this idea rigorous.

The flexibility of the methods presented here lies in the fact that all suggested tests, in Corollaries 27–30, can be combined in an attempt to prove existence and/or uniqueness in (2).

Proposition 24. *Let* P_1 *and* P° *be* Λ*-invariant* P*-parts. Suppose* P_1 *intersects* D *and*:

There exists a $\|\cdot\|$*-neighborhood* W *of* P_1 *such that for every* $n \geqslant 1$ *and every* $\{\mathcal{C}, \tilde{\Lambda}(\mathcal{C}), \ldots, \tilde{\Lambda}^n(\mathcal{C})\} \subset W \cap \mathrm{P}^\circ \cap \mathrm{D}$ *there exist* $f \in \operatorname{im} \mathrm{P}_1 \setminus \{0\}$ *and* $g \in \ker \mathrm{P}_1 \setminus \ker \mathrm{P}^\circ$ *with*

$$\frac{\Lambda^n(\mathcal{C})(g)}{\Lambda^n(\mathcal{C})(f)} \geqslant \frac{\mathcal{C}(g)}{\mathcal{C}(f)}. \tag{20}$$

Let $\mathcal{A} \in \mathrm{P}^\circ \cap \mathrm{D}$ *and consider its orbit* $(\tilde{\Lambda}^m(\mathcal{A}))_m$*. There are two possibilities*:

(i) *Either* W *is a* $\|\cdot\|$*-neighborhood of* $\overline{\mathrm{P}_1} \cap \mathrm{D} \cap S$*, then* $(\tilde{\Lambda}^m(\mathcal{A}))_m$ *does not accumulate at* P_1,

(ii) *or* W *is no* $\|\cdot\|$*-neighborhood of* $\overline{\mathrm{P}_1} \cap \mathrm{D} \cap S$*, then for every* $(\tilde{\Lambda}^m(\mathcal{A}))_m$ *accumulating at* P_1 *there exists another* P*-part* P_2 *satisfying*:

- $(\tilde{\Lambda}^m(\mathcal{A}))_m$ *also accumulates at* P_2,

- $\mathrm{P}_2 < \mathrm{P}_1$ *and*
- P_2 *contains an eigenvector*.

Proof. Let P_1^+ denote the set of all forms in P whose kernels are subsets of $\ker \mathrm{P}_1$. For $\mathcal{D} \in \mathrm{P}_1^+$ define

$$d_1(\mathcal{D}) := \frac{\sup\{\mathcal{D}(h) | h \in \ker \mathrm{P}_1, \|h\|_2 = 1\}}{\inf\{\mathcal{D}(h) | h \in \operatorname{im} \mathrm{P}_1, \|h\|_2 = 1\}}.$$

Then d_1 vanishes on P_1 and is positive on all forms in P_1^+ whose kernels are strict subsets of $\ker \mathrm{P}_1$.

(a) Suppose the orbit enters W at $\tilde{\Lambda}^m(\mathcal{A}) =: \mathcal{C} \in W$ and stays in W until $\tilde{\Lambda}^{m+n}(\mathcal{A}) \in W$ for some $n \geqslant 1$. Then

$$d_1(\tilde{\Lambda}^n(\mathcal{C})) \geqslant \frac{\Lambda^n(\mathcal{C})(g)}{\Lambda^n(\mathcal{C})(f)} \geqslant \frac{\mathcal{C}(g)}{\mathcal{C}(f)} > 0.$$

So the orbit cannot accumulate at P_1 as long as it stays in W.

(b) Suppose W is a $\|\cdot\|$-neighborhood of $\overline{\mathrm{P}_1} \cap \mathrm{D} \cap S$. The set $\overline{\mathrm{P}_1}$ consists of all P-parts $\mathrm{P}_2 \leqslant \mathrm{P}_1$ and 0. By Lemma 6, W is infinitely h-deep for all these P-parts. Furthermore, every single step of the orbit is bounded above by $h(\mathcal{A}, \Lambda(\mathcal{A})) =: r < \infty$. Thus $(\mathrm{P}^\circ \setminus W)_r$ does not accumulate at $\overline{\mathrm{P}_1}$. This means d_1 is uniformly positive on $(\mathrm{P}^\circ \setminus W)_r$ because $\overline{\mathrm{P}_1} \cup \mathrm{P}_1^+$ is a $\|\cdot\|$-neighborhood of $\overline{\mathrm{P}_1}$. According to (a) this implies $\omega(\mathcal{A}) \cap \mathrm{P}_1 = \emptyset$.

(c) Suppose $\omega(\mathcal{A}) \cap \mathrm{P}_1 \neq \emptyset$. As a consequence of (a) the orbit has to oscillate around $\partial W \cap \mathrm{P}^\circ$ such that there is a sequence $(k_m)_m \subset \mathrm{N}$ of successive reentries into W satisfying:

$$\tilde{\Lambda}^{k_m}(\mathcal{A}) \in W, \tilde{\Lambda}^{k_m - 1}(\mathcal{A}) \in \mathrm{P}^\circ \setminus W \text{ for } m \geqslant 1,$$
$$\lim_{m \to \infty} d_1(\tilde{\Lambda}^{k_m}(\mathcal{A})) = 0. \tag{21}$$

Since $(\mathrm{B}, \|\cdot\|)$ is locally compact, we may suppose without loss of generality that $(\tilde{\Lambda}^{k_m - 1}(\mathcal{A}))_m$ converges. As a consequence of $\Lambda \in C(\mathrm{D})$, $(\tilde{\Lambda}^{k_m}(\mathcal{A}))_m$ also converges. We know that every single step of the orbit is bounded above by $r < \infty$ and that W is infinitely h-deep. Thus the limit of $(\tilde{\Lambda}^{k_m}(\mathcal{A}))_m$ is an element of $\overline{W} \setminus \mathrm{P}_1$. Using (21) we deduce that $\omega(\mathcal{A})$ intersects the set $(\overline{W} \setminus \mathrm{P}_1) \cap \partial \mathrm{P}$.

(d) Suppose $\omega(\mathcal{A}) \cap \mathrm{P}_1 \neq \emptyset$ and W is no $\|\cdot\|$-neighborhood of $\overline{\mathrm{P}_1} \cap \mathrm{D} \cap S$. Using (c) for this W, we see that $\omega(\mathcal{A})$ intersects $\overline{\mathrm{P}_1} \setminus \mathrm{P}_1$ provided it is nonvoid. It is void if and only if P_1 is a single ray. But then W is also a neighborhood of $\overline{\mathrm{P}_1} \setminus \{0\} \cap S$ which is a contradiction. So we continue with the case $\overline{\mathrm{P}_1} \setminus \mathrm{P}_1 \neq \emptyset$. The P-part P_2' of a point in $\omega(\mathcal{A}) \cap (\overline{\mathrm{P}_1} \setminus \mathrm{P}_1)$ is Λ-invariant by Lemma 5. Using Proposition 15 for P_2'

we find an eigenvector in $\overline{\mathrm{P}_2'}$. Its P-part is the desired P_2, because $\mathrm{P}_2 \leqslant \mathrm{P}_2' < \mathrm{P}_1$, as a consequence of $\mathrm{P}_2' \subset \overline{\mathrm{P}_1} \setminus \mathrm{P}_1$.

Statement (ii) of Proposition 24 is the way how we can avoid Sabot's Assumption H. The technical proposition allows to prove a variety of existence results. They all prove the non attraction statement (20) but in different, more practical ways.

Theorem 25 (*Existence*). *Suppose* D° *is* Λ*-invariant. For every* Λ*-invariant* D*-part* $\mathrm{D}_i \subset \partial\mathrm{P}$ *containing an eigenvector* $\mathcal{E}^i$ *with eigenvalue* λ_i *consider the* P*-part* P_i *containing it. When* (20) *is satisfied for all these* P_i*, then there exists an eigenvector in* $\mathrm{P}^\circ \cap \mathrm{D}$.

Proof. (a) When D° is Λ-invariant, then $\mathrm{P}^\circ \supset \mathrm{D}^\circ$ also is. There are only finitely many D-parts. Denote the specific D-parts in the assumption by $\mathrm{D}_1, \ldots, \mathrm{D}_l$. We can apply Proposition 24 to every P_i. Look at a minimal P_i. It contains no accumulation point, because there is no smaller P-part to accumulate at. This argument works for all minimal P-parts. Take a second smallest P-part now. It also contains no accumulation point, since there is no smaller P-part left at which the orbit might accumulate. Repeat this argument until every P_i turns out to be disjoint from $\omega(\mathcal{A})$ for one (every) $\mathcal{A} \in \mathrm{P}^\circ \cap \mathrm{D}$.

(b) According to Proposition 15 an h-unbounded orbit $(\tilde{\Lambda}^n(\mathcal{A}))_n$ started in $\mathcal{A} \in \mathrm{P}^\circ \cap \mathrm{D}$ accumulates at a $\mathcal{B} \in \mathrm{D} \cap \partial\mathrm{P}$ for which there exists an eigenvector $\mathcal{E} \in \mathrm{D} \cap \partial\mathrm{P}$ with $\mathrm{P}_{\mathcal{B}} = \mathrm{P}_{\mathcal{E}}$. Obviously $\mathrm{D}_{\mathcal{E}}$ and $\mathrm{P}_{\mathcal{E}}$ are Λ-invariant. By (a), $\omega(\mathcal{A}) \cap \mathrm{P}_{\mathcal{E}} = \emptyset$. This is a contradiction. Thus the orbit must be h-bounded. Lemma 14 guarantees the existence of an eigenvector $\mathcal{E} \in \mathrm{P}^\circ$. Since the orbit was started in D it stays there and $\mathcal{E} \in \mathrm{D}$. □

There are indeed finitely ramified fractals which have an eigenvector in $\mathrm{P}^\circ \cap \mathrm{D}$ but not in D°. An example is the modified Sierpinski triangle in [30] which is actually a tree.

Our next task is to show how existence and uniqueness can be proved simultaneously when inequality (20) becomes strict. The underlying idea is to find additional eigenvectors in S_r, as defined below, when there are linear independent eigenvectors in $\mathrm{P}^\circ \cap \mathrm{D}$. The observation is due to Barlow, [3, Section 5], but it was turned into a powerful tool by Sabot [32, Section 5.4].

Theorem 26 (*Existence and uniqueness*). *Suppose the assumptions of Theorem* 25 *are satisfied with strict inequalities in* (20). *When there exists only one minimal* Λ*-invariant* D*-part in* P°*, then the eigenvector of Theorem* 25 *is unique* (*up to positive multiples*).

Proof. This is an extension of [32, Lemma 5.13] to graph-directed fractals. By Theorem 25 there exists an eigenvector $\mathcal{E} \in \mathrm{P}^\circ \cap \mathrm{D}$. Denote its D-part by D_0.

(a) Suppose there exists another eigenvector $\mathcal{E}' \in \mathrm{D}_0$ linear independent of $\mathcal{E}$. Since both eigenvectors lie in P°, there eigenvalues must coincide, say they equal λ. Define $\mathcal{B} := \mathcal{E}' - \inf(\mathcal{E}'/\mathcal{E})\mathcal{E}$. Obviously $h(\mathcal{E}, \mathcal{E}') = h(\Lambda(\mathcal{E}), \Lambda(\mathcal{E}'))$ and consequently

$\inf[\mathcal{E}'/\mathcal{E}] = \inf[\Lambda(\mathcal{E}')/\Lambda(\mathcal{E})]$. By (4) this implies

$$\Lambda(\mathcal{B} + \inf[\mathcal{E}'/\mathcal{E}] \cdot \mathcal{E})(f) = \Lambda(\mathcal{B})(f) + \inf[\mathcal{E}'/\mathcal{E}] \cdot \Lambda(\mathcal{E})(f)$$

for all $f \in \ker \mathcal{B}$. Thus $\Lambda(\cdot)(f)$ is linear on the plane spanned by $\mathcal{B}$ and $\mathcal{E}$ intersected with P according to Lemma 21. Additionally, (4) shows $\Lambda(\mathcal{B})(f) = 0$. This results in $\Lambda(\mathcal{C})(f) = 0$ for all $\mathcal{C} \in \overline{\mathrm{P}_{\mathcal{B}}}$ because Λ is P-increasing. Using an $\alpha > 0$ such that $\mathcal{C} \leqslant \alpha\mathcal{B}$ this implies

$$\Lambda(\mathcal{C} + \mathcal{E})(f) \leqslant \Lambda(\alpha\mathcal{B} + \mathcal{E})(f) = \Lambda(\mathcal{E})(f).$$

The reversed inequality is always true, since Λ is P-increasing. We arrive at

$$\Lambda(\mathcal{E} + \mathcal{C})(f) = \Lambda(\mathcal{E})(f) = \lambda\mathcal{E}(f) \tag{22}$$

for all $\mathcal{C} \in \overline{\mathrm{P}_{\mathcal{B}}}$ and all $f \in \ker \mathcal{B}$. Analogously, using the equality of $\sup[\mathcal{E}/\mathcal{E}']$ and $\sup[\Lambda(\mathcal{E})/\Lambda(\mathcal{E}')]$ together with $\mathcal{B}' := \sup[\mathcal{E}/\mathcal{E}']\mathcal{E}' - \mathcal{E}$ we get

$$\Lambda(r\mathcal{E} - \mathcal{D})(g) = \Lambda(r\mathcal{E})(g) = \lambda r\mathcal{E}(g) \tag{23}$$

for all $r \geqslant 0$, all $g \in \ker \mathcal{B}'$ and all $\mathcal{D} \in \overline{\mathrm{P}_{\mathcal{B}'}}$ such that $r\mathcal{E} - \mathcal{D} \in \mathrm{P}$. For $r \geqslant 1$ the Eqs. (22) and (23) show that

$$\mathrm{S}_r := \{\mathcal{A} \in \overline{\mathrm{D}_0} | \mathcal{E} \leqslant \mathcal{A} \leqslant r\mathcal{E},\, \mathcal{A} - \mathcal{E} \in \overline{\mathrm{P}_{\mathcal{B}}},\, r\mathcal{E} - \mathcal{A} \in \overline{\mathrm{P}_{\mathcal{B}'}}\}$$

is Λ/λ-invariant. It is also closed, convex and bounded because of its definition and the convexity of $\mathrm{P}_{\mathcal{B}}$, $\mathrm{P}_{\mathcal{B}'}$. Brouwer's fixed point theorem gives a fixed point of Λ/λ provided $\mathrm{S}_r \neq \emptyset$. Because of $\mathcal{E}' \in \mathrm{S}_1$, the set $\mathcal{E} + \overline{\mathrm{P}_{\mathcal{B}}}$ intersects $\overline{\mathrm{D}_0}$. Thus $\mathrm{S}_r \neq \emptyset$ if and only if $\overline{\mathrm{D}_0} \cap (r\mathcal{E} - \overline{\mathrm{P}_{\mathcal{B}'}}) \neq \emptyset$. Since $\overline{\mathrm{D}_0}$ and $\overline{\mathrm{P}_{\mathcal{B}'}}$ are closed convex cones, there is a critical value $r' \in [1, +\infty]$ such that $\mathrm{S}_r \neq \emptyset$ if and only if $1 \leqslant r \leqslant r'$. Since $h(\mathcal{E}, \mathcal{C}) = \ln r$, $r' = +\infty$ means that we can produce a sequence of eigenvectors in $\overline{\mathrm{D}_0}$ accumulating at $\partial\mathrm{P} \cap \overline{\mathrm{D}_0}$. This contradicts the strict inequality in (20). So $r' < +\infty$ and we have found an eigenvalue in $(\overline{\mathrm{D}_0} \setminus \mathrm{D}_0) \cap \mathrm{P}^\circ$. This means there exists a D-part $\mathrm{D}_1 < \mathrm{D}_0$ in P° which contains an eigenvector with eigenvalue λ.

(b) Let D_1 be the minimal D-part in P° containing an eigenvalue $\mathcal{E}^1$. Suppose there exists $\mathcal{E}^2 \in \mathrm{D}_1$ linear independent of $\mathcal{E}^1$ which is also Λ/λ-fixed. Then (a) proves the existence of a Λ/λ-fixed point $\mathcal{E}^3 \in (\overline{\mathrm{D}_1} \setminus \mathrm{D}_1) \cap \mathrm{P}^\circ$. Thus the D-part of $\mathcal{E}^3$ is strictly smaller than D_1. This is a contradiction.

(c) We will prove that $\mathcal{E} =: \mathcal{E}^0$ is unique in D_0. Suppose not, then the minimal D-part D_1 must be contained in $\overline{\mathrm{D}_0} \setminus \mathrm{D}_0$ so that (a) cannot cause a contradiction. Let $\mathcal{E}^1$ be the unique Λ/λ-fixed point in D_1. Define S_1 by $\mathcal{E}^0$ and $\mathcal{E}^1$. With (a) we produce another fixed point $\mathcal{E}^2 \in (\overline{\mathrm{D}_0} \setminus \mathrm{D}_0) \cap \mathrm{P}^\circ$. We know that it cannot be in D_1 because of (b) and $h(\mathcal{E}^1, \mathcal{E}^2) = \ln r$. So it must be an element of a D-part $\mathrm{D}_1 < \mathrm{D}_2 < \mathrm{D}_0$ by (a) and the minimality of D_1. Now we repeat the argument with $\mathcal{E}^1$ and $\mathcal{E}^2$ to produce another $\tilde{\Lambda}$-fixed point in a D-part strictly between $\mathrm{D}_{\mathcal{E}^2}$ and $\mathrm{D}_{\mathcal{E}^1}$. Since there are only finitely many D-parts, we end up with a contradiction after finitely many repetitions of the argument. □

The function test below is designed to work in cases of ambiguous eigenvectors in $\mathrm{P}^\circ \cap \mathrm{D}$. In the one-dimensional dynamics of differentiable maps on the interval the derivative condition is replaced by a condition on the function itself when the derivative equals one [31, p. 22]. Remark 36 explains why ambiguous eigenvectors cause a similar effect in our multidimensional situation. Therefore we adopt the one-dimensional strategy. The assumptions of the function test are quite strong and possibly not necessary. The test can also be used in cases of uniqueness when its inequality is strict. The diamond fractal in [14] is an example it applies to but it fails on the Hany fractal.

Corollary 27 (*Function test*). *Suppose the* D-*parts* D° *and* $\mathrm{D}_1 \subset \partial \mathrm{P}$ *are* Λ-*invariant,* D_1 *contains an eigenvector and is contained in the* P-*part* P_1. *Let* W *be a* $\|\cdot\|$-*neighborhood of* P_1 *for which there exist* $f \in \operatorname{im} \mathrm{P}_1 \setminus \{0\}$ *and* $g \in \ker \mathrm{P}_1 \setminus \ker \mathrm{P}^\circ$ *with*

$$\frac{\Lambda(\mathcal{C})}{\mathcal{C}}(f) \leqslant \frac{\Lambda(\mathcal{C})}{\mathcal{C}}(g) \quad \text{for all } \mathcal{C} \in W \cap \mathrm{P}^\circ \cap \mathrm{D},$$

then (20) *holds for* P_1.

Proof. Suppose $\mathcal{C}, \tilde{\Lambda}(\mathcal{C}), \ldots, \tilde{\Lambda}^n(\mathcal{C})$ are contained in W. Then

$$\frac{\Lambda^n(\mathcal{C})}{\mathcal{C}}(h) = \frac{\Lambda(\mathcal{C})}{\mathcal{C}}(h) \cdot \frac{\Lambda^2(\mathcal{C})}{\Lambda(\mathcal{C})}(h) \cdots \frac{\Lambda^n(\mathcal{C})}{\Lambda^{n-1}(\mathcal{C})}(h)$$

for $h \notin \ker \mathcal{C}$. Now use $h \in \{f, g\}$ to prove (20). □

In Corollary 27 it is quite easy to estimate $\frac{\Lambda(\mathcal{C})}{\mathcal{C}}(g)$ from below for $\mathcal{C}$ in a convex set. The reason is Lemma 12.

The next corollary is an improved version of Sabot's existence result, [32, Theorem 5.1(ii)], and therefore called Sabot's test. The main differences between the original and our version are: Sabot's very restrictive Assumption H has been removed completely; the spectral formulation of the assumptions allows to apply the Corollary recursively and to use the classical eigenvalue estimates of Section 3.2; the theorem applies to finitely ramified graph-directed constructions, a much bigger class than the previously considered single component fractals. The test applies to the Hany fractal but not to the Vicsek set in Section 8.1.

Corollary 28 (*Sabot's test*). *Suppose the* D-*parts* D° *and* $\mathrm{D}_1 \subset \partial \mathrm{P}$ *are* Λ-*invariant,* D_1 *contains an eigenvector* $\mathcal{E}$ *with eigenvalue* λ *and is contained in the* P-*part* P_1.

When the map $\Lambda(\infty\mathcal{E}+\cdot)$ has an eigenvalue $\gamma>\lambda$, then the inequality in (20) *is strict for* P_1.

Proof. There exists $0<\varepsilon<\frac{\gamma-\lambda}{\gamma+\lambda}$. Let U be the neighborhood in Proposition 23 for this ε. By our assumptions the orbit $(\tilde{\Lambda}^n(\mathcal{A}))_n\subset \mathrm{P}^\circ\cap\mathrm{D}$. Suppose it enters U, say in $\mathcal{C}\in\mathrm{P}^\circ\cap\mathrm{D}$.

Let $k\in\mathrm{N}$. Because of λ and Corollary 13, there exists $g\in\operatorname{im}\mathrm{P}_1\setminus\{0\}$ such that

$$\Lambda^k\circ\Pi_1(\mathcal{C})(g)\leqslant\lambda_i^k\cdot\Pi_1(\mathcal{C})(g).$$

The eigenvalue γ has an eigenvector in $\infty\mathcal{E}+\mathrm{D}$. According to Proposition 11, $\mathrm{cwi}_k(\infty\mathcal{E}+(\mathrm{D}\cap\mathrm{P}^\circ))$ contains a real not less than γ. Hence there exists $f\in\ker\mathrm{P}_1\setminus\ker\mathrm{P}^\circ$ such that

$$\Lambda^k\circ\Pi_\infty(\mathcal{C})(f)\geqslant\gamma^k\cdot\Pi_\infty(\mathcal{C})(f).$$

Now suppose $\{\tilde{\Lambda}^n(\mathcal{C})|1\leqslant n\leqslant k\}\subset U$. By our choice of U and (19),

$$\Pi_1\circ\Lambda^k(\mathcal{C})(g)\leqslant(1+\varepsilon)^k\lambda^k\cdot\Pi_1(\mathcal{C})(g),$$
$$\Pi_\infty\circ\Lambda^k(\mathcal{C})(f)\geqslant(1-\varepsilon)^k\gamma^k\cdot\Pi_\infty(\mathcal{C})(f).$$

The choice of ε guarantees the existence of an $\eta>0$ such that

$$\frac{\Pi_\infty\circ\Lambda^k(\mathcal{C})(f)}{\Pi_1\circ\Lambda^k(\mathcal{C})(g)}\geqslant(1+\eta)^k\cdot\frac{\Pi_\infty(\mathcal{C})(f)}{\Pi_1(\mathcal{C})(g)}.$$

Since $f\in\ker\mathrm{P}_1$, we have $\Pi_\infty(\mathcal{D})(f)=\mathcal{D}(f)$ for every $\mathcal{D}\in\mathrm{P}$. Since Π_1 is continuous and $g\in\operatorname{im}\mathrm{P}_1$, there exists $0<\alpha<1+\eta$ such that $\frac{1}{\alpha}\mathcal{D}(g)\leqslant\Pi_1(\mathcal{D})(g)\leqslant\alpha\mathcal{D}(g)$ for all $\mathcal{D}$ in a small neighborhood of P_1. We intersect it with U to form W. On W we have

$$\frac{\Lambda^k(\mathcal{C})(f)}{\Lambda^k(\mathcal{C})(g)}\geqslant\frac{(1+\eta)^k}{\alpha^2}\cdot\frac{\mathcal{C}(f)}{\mathcal{C}(g)}>\frac{\mathcal{C}(f)}{\mathcal{C}(g)}. \tag{24}$$

Now Theorem 25 applies with a strict inequality in (20). $\square$

The next Corollary is derived from Sabot's test. It applies to the diamond fractal [25] but not to the Hany fractal (because $\gamma=\lambda$).

Corollary 29 (*Eigenvalue test*). *Suppose the* D*-parts* D° *and* $\mathrm{D}_1\subset\partial\mathrm{P}$ *are* Λ*-invariant. Let* D_1 *contain an eigenvector* $\mathcal{E}$ *with eigenvalue* λ *and denote its* P*-part by* P_1. *When there exists an* $\mathcal{A}\in\mathrm{D}\setminus\{0\}$ *such that* $\Lambda(\mathcal{A})\geqslant\gamma\mathcal{A}$ *for some* $\gamma>\lambda$, *then the inequality in* (20) *is strict for* P_1.

Proof. Corollary 11 and $\gamma > \lambda$ imply $\mathrm{P}_1 < \mathrm{P}_{\mathcal{A}}$. Thus $\infty\mathcal{E} + \mathcal{A}$ is nonzero in D_∞. Furthermore, $\Lambda(\infty\mathcal{E} + \mathcal{A}) \geqslant \gamma(\infty\mathcal{E} + \mathcal{A})$ according to Lemma 20. Hence $\Lambda(\infty\mathcal{E} + \cdot)$ has an eigenvalue not less than γ by Lemma 8. Now apply Sabot's test. $\square$

One can also prove the eigenvalue test directly via (3) like in [25, Theorem 1]. The homogenization results in [30,24] show that a numerical iteration of the Schur complement formula approximates the leading eigenvector of Λ quite rapidly. The resulting data, although imprecise, can be used to estimate the γ of the eigenvalue test. A consequence of the eigenvalue test is Nussbaum's test.

Corollary 30 (*Nussbaum's test [29, Theorem 3.4]*). *Suppose the* D*-parts* D° *and* $\mathrm{D}_1 \subset \partial\mathrm{P}$ *are* Λ*-invariant. Let* D_1 *contain an eigenvector* $\mathcal{E}$ *with eigenvalue* λ *and denote its* P*-part by* P_1. *When there exists an* $\mathcal{A} \in \mathrm{D} \setminus \{0\}$ *such that* $\frac{\partial}{\partial \mathcal{A}} \Lambda_{\mathcal{E}} \geqslant \gamma\mathcal{A}$ *for some* $\gamma > \lambda$, *then the inequality in* (20) *is strict for* P_1.

Proof. Consider $\mathcal{E} + \varepsilon\mathcal{A}$ for $\varepsilon > 0$. Because of Lemma 22(i) we can find an $\varepsilon > 0$ such that $\Lambda(\mathcal{E} + \varepsilon\mathcal{A}) > \lambda(\mathcal{E} + \varepsilon\mathcal{A})$. Thus D contains an eigenvector with eigenvalue strictly bigger than λ by Lemma 8. Now apply Corollary 29. $\square$

The proof shows that Nussbaum's test is a way to verify the assumptions of the eigenvalue test. Thus it does not work for the Hany fractal.

Finally, lets us try the methods of this section on the Hany fractal. We have identified the reducible, Λ-invariant D-parts to be D_1 and D_3. The eigenvalue of D_1 is $\frac{1}{3}$ and we have deduced from (6) that all other eigenvalues also equal $\frac{1}{3}$. Since $\mathrm{D}_1 \leqslant \mathrm{D}_3$, we can try Theorem 26 on D_3 to find an eigenvalue. The graph $\Gamma((\infty\mathrm{D}_1 + \mathrm{D}_3)_1)$ is shown in the third column of Fig. 2 when the solid black lines (switched off by $d_3 = 0$) are removed. The shorted triangle has two vertices $\{v_1 = v_3, v_2\}$, since d_1 connects v_1 to v_3. The conductance between the two vertices is $\frac{2d_2d_4}{2d_2+d_4}$. The square still has four vertices arranged in two connected components $\{v_4, v_6\}$ and $\{v_5, v_7\}$ as we can see from (5). One connected component is mapped onto the other by $\mathfrak{G}$, so it suffices to calculate the conductances in one of them; it is $\frac{d_2d_4}{d_2+d_4}$. The corresponding conductances in the initial graph are $2d_2$ on the shorted triangle and d_4 on each connected component of the shorted square. So $\Lambda(\infty\mathrm{D}_1 + \cdot)$ scales on D_3 with the two extremal values

$$\frac{d_4}{2d_2 + d_4} \text{ and } \frac{d_2}{d_2 + d_4}. \tag{25}$$

For $d_2, d_4 > 0$ they are equal if and only if $d_4 = \sqrt{2} \cdot d_2$. The corresponding eigenvalue is $\frac{1}{1+\sqrt{2}}$. Since this is strictly bigger than $\frac{1}{3}$, we have a unique eigenvector in D_3. The corresponding eigenvector has the conductances $(1, 2, 0, 4)$ according to [24, Theorem 4] but we will not need this information. The function test fails in this situation because there is no single $g \in \ker \mathrm{D}_1 \setminus \ker \mathrm{P}^\circ$ meeting the requirements of the test!

The short circuited graph $\Gamma(\infty\mathrm{D}_3 + \mathrm{D}^\circ)$ is shown in the fourth column of Fig. 2. Only one vertex remains of the shorted triangle, so it is irrelevant for scaling questions. The shorted square has two vertices $v_4 = v_6$ and $v_5 = v_7$. The conductance between them is $16d_3$. The corresponding value for the initial square is $4d_3$. Hence $\Lambda(\infty\mathrm{D}_3 + \cdot)$ scales on D° with 4. This is much bigger than $\frac{1}{3}$. It remains to calculate the scaling of $\Lambda(\infty\mathrm{D}_1 + \cdot)$ in D°. Since $\infty\mathrm{D}_1 + \mathrm{D}_3$ is a subset of $\infty\mathrm{D}_1 + \mathrm{D}$, we already have an eigenvalue $\frac{1}{1+\sqrt{2}} > \frac{1}{3}$. Thus Sabot's test is positive and we have proved existence. Because of Theorem 26 the corresponding eigenvector is unique.

Remark 31. The repeated use of $\frac{1}{1+\sqrt{2}}$ in the above argument on the Hany fractal can be turned into a theorem: Let $\mathrm{D}_1, \mathrm{D}_2$ be two Λ-invariant D-parts lying in P-parts $\mathrm{P}_1 < \mathrm{P}_2$. Then P_1 is a Λ-invariant $\overline{\mathrm{P}_2}$-part of $\overline{\mathrm{P}_2}$. When Sabot's test proves that P_1 is weakly repellent in $\overline{\mathrm{P}_2}$, then P_1 is also weakly repellent in P. The proof is the same as in the Hany fractal case.

7. Nonexistence

Suppose P° is Λ-invariant and $\partial\mathrm{P} \cap \mathrm{D}$ contains eigenvectors. Since D has only finitely many D-parts, there are only finitely many P-parts $\mathrm{P}_1, \ldots, \mathrm{P}_k$ with eigenvector $\mathcal{B}_i \in \mathrm{P}_i \cap \mathrm{D}$ and eigenvalue λ_i. We can prove the analog of a well known linear result: An orbit started in the interior of the cone is attracted by the leading eigenvectors (defined to have leading eigenvalues).

Corollary 32. *Let* P° be *Λ-invariant and* $\mathcal{A} \in \mathrm{P}^\circ \cap \mathrm{D}$ *such that* $(\Lambda^n(\mathcal{A}))_n$ *is h-unbounded. Then* $(\tilde{\Lambda}_n(\mathcal{A}))_n$ *accumulates only at points* $\mathcal{B} \in \mathrm{D}$ *such that* $\mathrm{P}_{\mathcal{B}} < \mathrm{P}^\circ$ *and* $\mathrm{P}_{\mathcal{B}} \cap \mathrm{D}$ *has the eigenvalue* $\max\{\lambda_i | 1 \leqslant i \leqslant k\}$.

Proof. According to the eigenvalue test every P_i whose eigenvalue is not maximal is nonattracting. □

Corollary 32 allows to restrict our existence tests to P-parts with leading eigenvalues! Nussbaum's test detects these parts and Remark 31 tells us how to deal with them. For $1 \leqslant i \leqslant k$ set $\gamma_i := \min \mathrm{cwi}(\infty\mathcal{B}_i + \mathrm{P}^\circ \cap \mathrm{D})$.

Proposition 33. *Suppose* P° *is Λ-invariant but* $\mathrm{P}^\circ \cap \mathrm{D}$ *contains no eigenvector. Then* $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ *equals* $\min\{\gamma_i | 1 \leqslant i \leqslant l\}$.

Proof. According to Lemma 20 we have $\gamma_i \geqslant \min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ for all $1 \leqslant i \leqslant k$. So it suffices to find a γ_i assuming $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$. Since $\mathrm{P}^\circ \cap \mathrm{D}$ contains no eigenvector, there must be a nonrepelling P_i. Denote the leading eigenvalue of $\Lambda(\infty\mathcal{B}_i + \cdot)$ on D by ρ_i. Then Sabot's test and Corollary 13 imply

$$\lambda_i \geqslant \rho_i = \max \mathrm{cwi}(\infty\mathcal{B}_i + \mathrm{P}^\circ \cap \mathrm{D}) \geqslant \gamma_i. \tag{26}$$

Let U be the union of all nonrepelling P_i, $1 \leqslant i \leqslant k$. Then every $\|\cdot\|$-neighborhood V of U has the property $\min \mathrm{cwi}(\mathrm{P}^\circ \cap V) = \min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$. Indeed, suppose the contrary. Then we can find an $\mathcal{A} \in \mathrm{P}^\circ \cap \mathrm{D}$ with $\inf[\Lambda(\mathcal{A})/\mathcal{A}] > \min \mathrm{cwi}(\mathrm{P}^\circ \cap V)$. Its orbit cannot accumulate at U by Proposition 3. Thus its orbit is bounded and there exists an eigenvalue in $\mathrm{P}^\circ \cap \mathrm{D}$ because of Lemma 14, a contradiction. Since $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ is assumed in every neighborhood of U, we can recover this value when we calculate $[\Lambda(\cdot)/\cdot]$ at U in the sense of (14) and (15). Let us do this at each $\mathcal{B}_i$ which is contained in U. Here (14) gives us $\{\lambda_i\}$. By Corollary 13 this is the biggest lower bound of all $[\Lambda(\cdot)/\cdot]$ on $\mathrm{P}_{\mathcal{B}_i}$. The values of (15) are $[\Lambda(\infty\mathcal{B}_i + \mathcal{D})/(\infty\mathcal{B}_i + \mathcal{D}]$ depending on the direction $\mathcal{D} \in \mathrm{P}^\circ \cap \mathrm{D}$. According to (26) the biggest lower bound is $\min \mathrm{cwi}(\infty\mathcal{B}_i + \mathrm{P}^\circ \cap \mathrm{D}) = \gamma_i$. Now taking the smallest γ_i with $\mathcal{B}_i \in U$ we recover $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$. $\square$

The value $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ was first used in [5, Section 4]. Proposition 33 gives a spectral interpretation of it when every $\Lambda(\infty\mathcal{B}_i + \cdot)$ has an eigenvalue in $\mathrm{P}^\circ \cap \mathrm{D}$. Then $\min \mathrm{cwi}(\mathrm{P}^\circ \cap \mathrm{D})$ equals the minimum of all these eigenvalues. In general the minimum of all these eigenvalues is only an upper bound because of Corollary 13.

Remark 34 (*Nonexistence, Sabot [32, Theorem 5.1(i)]*). As a consequence of Corollary 13 and Proposition 33 we have: Suppose P° is Λ-invariant and $\partial\mathrm{P} \cap \mathrm{D}$ contains eigenvectors. Define λ_i, γ_i and k as above. When

$$\max\{\lambda_i | 1 \leqslant i \leqslant k\} > \min\{\gamma_i | 1 \leqslant i \leqslant k\},$$

then $\mathrm{P}^\circ \cap \mathrm{D}$ contains no eigenvector.

Sabot's test and Remark 34 tell us what happens when $\lambda_i \neq \gamma_i$ in the corresponding statements. It remains to consider the case of equality. It typically has to be considered when all coupling weights causing existence have to be calculated like in Section 8.2. For $\mathcal{D} \in \mathrm{P}^\circ \cap \mathrm{D}$ set

$$E_\lambda(\mathcal{D}) := \{f \in \ker\mathcal{B} \setminus \ker \mathrm{P}^\circ \ | \Lambda(\infty\mathcal{B} + \mathcal{D})(f) = \lambda(\infty\mathcal{B} + \mathcal{D})(f)\}.$$

Theorem 25 shows that existence now depends on the scaling of suitable $f \in E_\lambda(\mathcal{D})$ near $\mathcal{B}$. One possible way to guarantee nonexistence is the following nonlinearity argument.

Theorem 35. *Suppose* P° *is* Λ*-invariant,* $\partial\mathrm{P} \cap \mathrm{D}$ *contains an eigenvector* $\mathcal{B}$ *with eigenvalue* $\lambda > 0$. *When for every* $\mathcal{D} \in \mathrm{P}^\circ \cap \mathrm{D}$ *there exists an* $f \in E_\lambda(\mathcal{D})$ *such that* $H^{(\infty\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f$ *is not* $(\mathcal{B}+\mathcal{D})_1$*-harmonic on* $V_1 \setminus V_0$*, then* $\mathrm{P}^\circ \cap \mathrm{D}$ *contains no eigenvector.*

Proof. According to Lemma 22(ii) the map $\varepsilon \mapsto \Lambda(\mathcal{B}/\varepsilon + \mathcal{D})(f)$ is decreasing for every $f \in \ker\mathcal{B}$. Suppose it is constant. Then

$$H^{(\mathcal{B}+\mathcal{D})_1}_{V_1 \setminus V_0} f - H^{(\mathcal{B}/\varepsilon+\mathcal{D})_1}_{V_1 \setminus V_0} f$$

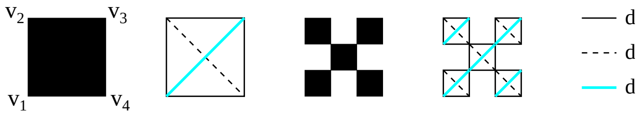

Fig. 3. The construction of the Vicsek fractal and its skeleton graphs.

is an element of $\ker(\mathcal{B}+\mathcal{D})_1$ by Proposition 2(ii). Since $\mathcal{B}+\mathcal{D} \in \mathrm{P}^\circ \cap \mathrm{D}$, the minimum principle shows that the two functions actually coincide. This is true for every $\varepsilon > 0$. Thus Proposition 18 implies that they also coincide with $H_{V_1 \setminus V_0}^{(\infty\mathcal{B}+\mathcal{D})_1} f$. This shows why our assumptions force $\varepsilon \mapsto \Lambda(\mathcal{B}/\varepsilon + \mathcal{D})(f)$ to be strictly decreasing. Hence $[\Lambda(\mathcal{B}+\varepsilon\mathcal{D})/(\mathcal{B}+\varepsilon\mathcal{D})]$ contains a real strictly less than λ. By Corollary 11 it also contains a real not less than λ. That is, it is a nontrivial interval for all $\varepsilon > 0$. Varying $\mathcal{D}$ we get the same result on $\mathcal{B}+\mathrm{P}^\circ$. Since the interval is constant on rays, the assertion follows. $\square$

8. Further examples

The following collection of examples helps to illustrate the methods derived so far or provides new existence results.

8.1. A case of nonuniqueness

The Vicsek fractal with reduced symmetry requirements is treated in [20]. Its construction and its initial graph with vertices $V_0 := \{v_1, \ldots, v_4\}$ and conductances $(d_1, d_2, d_3) \in \mathrm{R}_+^3$ are indicated in Fig. 3. We index its D-parts again by representative conductances $(d_1, d_2, d_3) \in \{0, 1\}^3$. The renormalization map of the Vicsek fractal is also denoted by Λ. Its action on D-parts can be found graphically via Lemma 16. The Λ-invariant D-parts and their connected components are:

$$\begin{aligned}
\mathrm{D}_{(1,1,1)} &= \mathrm{D}^\circ &&: \{v_1, v_2, v_3, v_4\},\\
\mathrm{D}_{(0,1,1)} &=: \mathrm{D}_1 &&: \{v_1, v_3\}, \{v_2, v_4\},\\
\mathrm{D}_{(0,1,0)} &=: \mathrm{D}_2 &&: \{v_1\}, \{v_2, v_4\}, \{v_4\},\\
\mathrm{D}_{(0,0,1)} &=: \mathrm{D}_3 &&: \{v_1, v_3\}, \{v_2\}, \{v_4\}.
\end{aligned} \tag{27}$$

We want to find an eigenvector in D°. It is not difficult to see that $\mathrm{D}_1, \mathrm{D}_2$ and D_3 consist of eigenvectors with eigenvalue $\lambda := \frac{1}{3}$. Like in (6) we can use the function $1_{v_2} - 1_{v_4}$ to prove that the eigenvalue of D° must also be $\frac{1}{3}$, when it exists at all.

It is shown in [20] that the eigenvectors in D° are $\{(d_1, d_2, d_3) \in \mathrm{R}_+^3 \,|\, d_1^2 = d_2 d_3 > 0\}$. Thus the above Vicsek fractal has two linear independent eigenvectors in $\mathrm{P}^\circ \cap \mathrm{D}$ with coinciding eigenvalue $\lambda = \frac{1}{3}$. Since D° is the only Λ-invariant D-part in P°, part (b) and (c) of the proof of Theorem 26 show that there exists a sequence of eigenvectors in $\mathrm{P}^\circ \cap \mathrm{D}$ accumulating at $\partial\mathrm{P} \cap \mathrm{D}$. Since Λ is continuous on D, this implies the existence

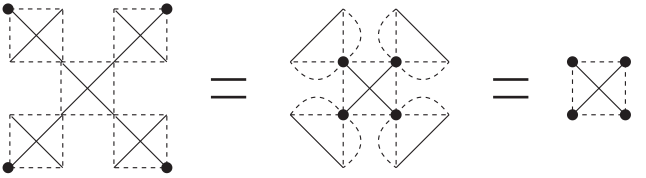

Fig. 4. Shorting the Vicsek fractal with conductance $+\infty$ or 1 on solid or dashed edges, respectively.

of an eigenvector $\mathcal{B} \in \partial \mathrm{P} \cap \mathrm{D}$ with eigenvalue λ. This is why Sabot's test never works in such cases of non-uniqueness!

Remark 36. Let $\mathcal{D} \in \mathrm{B}$ be a "direction" such that $\mathcal{B} + \mathcal{D}$ is an eigenvector in $\mathrm{P}^\circ \cap \mathrm{D}$. Then

$$\frac{\Lambda(\mathcal{B} + \varepsilon\mathcal{D})}{\mathcal{B} + \varepsilon\mathcal{D}}(g) \geqslant \lambda \text{ for } 0 \leqslant \varepsilon \leqslant 1 \text{ and } g \in \operatorname{im} \mathcal{B} \setminus \{0\}$$

according to Lemma 12. For $f \in \ker \mathcal{B} \setminus \ker \mathrm{P}^\circ$ the map $\varepsilon \mapsto \frac{\Lambda(\mathcal{B}+\varepsilon\mathcal{D})}{\mathcal{B}+\varepsilon\mathcal{D}}(f)$ is decreasing in ε. Hence it must be identically λ. So the best we can hope for, in order to avoid nonexistence, is that the corresponding map for g is also identically λ. This means $\Lambda(\cdot)(f)$ and $\Lambda(\cdot)(g)$ are linear on $\mathcal{B} + \mathrm{R}_+\infty$. One can, for example, use Gâteaux derivatives at $\partial \mathrm{P} \cap \mathrm{D}$ to detect such linearities and coinciding slopes. This is the "derivative 1" effect mentioned in Section 6.

Let us first look at D_1 and its eigenvalue $\frac{1}{3}$. We calculate λ_∞, the eigenvalue of $\Lambda(\infty\mathcal{B} + \cdot)$, graphically as shown in Fig. 4: $\Gamma(\Psi(\infty \mathrm{D}_1 + \mathrm{D}^\circ))$ is shown on the very left. The dashed edges are defined to have finite and positive conductance and the solid edges to have infinite conductance. The dots are the vertices of V_0. We first use Lemma 17 to identify the endpoints of those short circuited edges which end in a point of V_0. This does not change the energy according to (9). "Identical energies" is the meaning of the equality signs in the figure. Now we have dangling ends at every point of V_0. They do not contribute to the energy, so we eliminate them. The resulting graph to the very right is exactly $\Gamma(\infty \mathrm{D}_1 + \mathrm{D}^\circ)$. Hence $\lambda_\infty = 1$ and Sabot's test applies, because every eigenvalue must be strictly less than 1 [13, Lemma 3.1.10].

Remark 37. The effect, that Sabot's test applies by purely graphical arguments, is typical for nested fractals. Suppose $\mathrm{D}_1 \subset \partial \mathrm{P}$ is a Λ-invariant D-part. When one has to calculate $c_{\Lambda(\infty\mathcal{B}+\mathcal{D})}(x, y)$ for two points x, y in different connected components of $\Gamma(\mathcal{B})$ like in Fig. 4, then there always exists a cell $\psi_e(V_0)$, the central square in Fig. 4, such that the connected components of x and y both intersect the cell [32, Lemma 6.5]. Thus a shorting argument similar to the one in Fig. 4 is possible, that is, $c_{\Lambda(\infty\mathcal{B}+\mathcal{D})}(x, y) \geqslant c_{\infty\mathcal{B}+\mathcal{D}}(x, y)$. Since this argument works for every pair of different $\mathcal{B}$-components, we arrive at $\lambda_\infty \geqslant 1$. Again any Λ-eigenvalue is less than 1 by [13, Lemma 3.1.10], an application of the strong minimum principle.

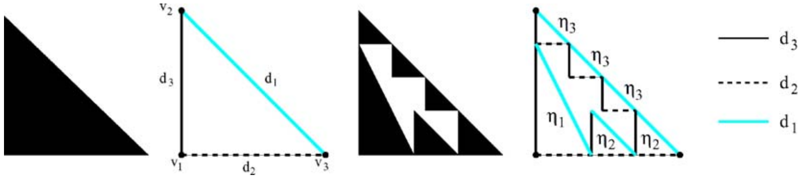

Fig. 5. The construction of the 123-gasket and its skeleton graphs.

Unfortunately, "Sabot's test" cannot be applied to D_2 and D_3 (because the eigenvectors in D° accumulate at these sets). So let us try the function test on D_2. The space $\operatorname{im} D_2$ is spanned by $g := 1_{v_2} - 1_{v_4}$ and

$$\frac{\Lambda(\mathcal{C})}{\mathcal{C}}(g) = \frac{1}{3} \text{ for all } \mathcal{C} \text{ with } \Gamma(D_2) \subset \Gamma(\mathcal{C}).$$

The kernel of D_1 consists of all functions which are constant on $\{v_2, v_3\}$ as can be seen from (27). So let us try, for example, $f := 1_{v_1} - 1_{v_3}$. We get the same scaling result as for g. Thus the function test is positive. For symmetry reasons an analogous argument works for D_3. This means, we have proved existence (without uniqueness) via Corollary 27 on the Vicsek set with reduced symmetry requirements.

8.2. The 123-gasket

In [10] so called *abc*-gaskets are considered as bad examples in terms of irreducible eigenvectors. They are modified Sierpinski gaskets with different numbers, $a, b, c \geqslant 1$, of triangles along the sides of the initial triangle. The 123-gasket can be seen in Fig. 5. The most symmetric version, the 111-gasket, is just the classical Sierpinski gasket. We will consider the 123-gasket and try to find positive and finite coupling weights such that an irreducible eigenvector exists in $D \cap P^\circ$. Let us call such a collection of coupling weights "good." A similar problem was solved for the Sierpinski gasket in [32, Section 5.2]. The answer was that the known good collection $(\eta_1, \eta_2, \eta_3) = (1, 1, 1)$ is surrounded by an open set of good weights (in $\mathbb{R}^3_+$).

The initial vertices are $V_0 := \{v_1, v_2, v_3\}$ and we arrange the three conductances $(d_1, d_2, d_3) \in \mathbb{R}^3_+$ and the three coupling weights $(\eta_1, \eta_2, \eta_3) \in (0, \infty)^3$ as in Fig. 5. We index the D-parts by representative conductances $(d_1, d_2, d_3) \in \{0, 1\}^3$. The action of Λ on D-parts does not depend on the coupling weights as long as they are all positive and finite. Therefore we have the same Λ-invariant D-parts as the Sierpinski gasket, namely, $D^\circ = D_{(1,1,1)}$, $D_1 := D_{(1,0,0)}$, $D_2 := D_{(0,1,0)}$ and $D_3 := D_{(0,0,1)}$. They are all located in different P-parts. The eigenvalue λ_i of D_i for $1 \leqslant i \leqslant 3$, is

$$\lambda_1 = \left[\frac{1}{\eta_2} + \frac{3}{\eta_3}\right]^{-1}, \quad \lambda_2 = \left[\frac{1}{\eta_1} + \frac{1}{\eta_3}\right]^{-1}, \quad \lambda_3 = \left[\frac{1}{\eta_1} + \frac{2}{\eta_2}\right]^{-1}. \tag{28}$$

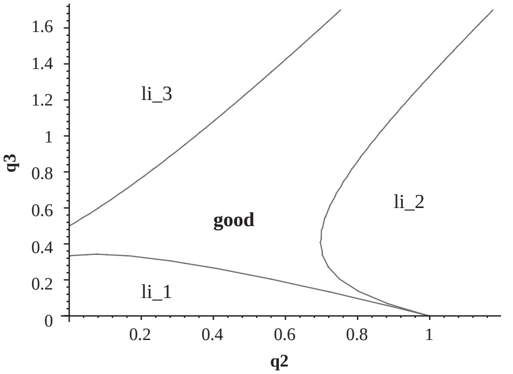

Fig. 6. The region of “good” coupling weights of the 123-gasket.

Shorting as in (9) we see that $\infty \mathrm{D}_i + \mathrm{D}^\circ$ consists of eigenvectors with eigenvalue γ_i, $1 \leqslant i \leqslant 3$,

$$\gamma_1 = \left[\frac{1}{\eta_1} + \left(\eta_3 + \frac{\eta_2}{2}\right)^{-1}\right]^{-1},$$

$$\gamma_2 = \left[\frac{1}{\eta_2} + \left(\frac{\eta_3}{3} + \left[\frac{1}{\eta_1} + \frac{1}{\eta_2}\right]^{-1}\right)^{-1}\right]^{-1},$$

$$\gamma_3 = \left[\frac{1}{\eta_3} + \left(\eta_1 + \left[\frac{2}{\eta_3} + \frac{1}{\eta_2}\right]^{-1}\right)^{-1}\right]^{-1}.$$

By Sabot’s test D_i is repellent if $\lambda_i < \gamma_i$. Since Λ is positively homogeneous, we can express this in the new variables $q_i := \eta_1/\eta_i$, $i \in \{2, 3\}$, by

$$2q_2^2 + 5q_2q_3 + 3q_3^2 - 2q_2 - q_3 > 0,$$

$$1 + q_3 + 3q_3^2 - q_2^2 - 5q_2q_3 > 0,$$

$$1 + 2q_2 + 2q_2^2 + 3q_2q_3 - q_3 - 2q_3^2 > 0.$$

The resulting region of “good” coupling weights is shown in Fig. 6. Here “li_i” denotes the line $\{\lambda_i = \gamma_i\}$ and the label indicates the “bad” side of the corresponding line which means the region of points not identified as “good” by Sabot’s test. Indeed Remark 34 identifies “bad” regions as regions of nonexistence and Theorem 35 proves nonexistence on the lines. The size of the good region shows that there is no qualitative difference between the 111- and the 123-gasket.

8.3. A self-similar graph without existence

The self-similar ladder graph in Fig. 7 is not the skeleton of a finitely ramified fractal. Nevertheless, it shows how non-existence can happen.

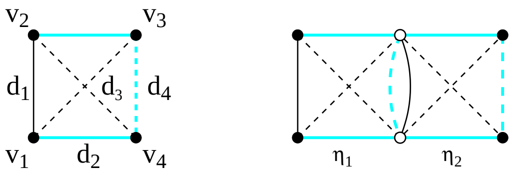

Fig. 7. The self-similar ladder graph.

This time $V_0 := \{v_1, \ldots, v_4\}$, D is four dimensional because of $d_1, \ldots, d_4 \in \mathrm{R}_+$ and two coupling weights $\eta_1, \eta_2 > 0$ are allowed. Indexed by representative conductances $(d_1, \ldots, d_4) \in \{0, 1\}^4$ the Λ-invariant D-parts are $\mathrm{D}_1 := \mathrm{D}_{(1,0,0,0)}$, $\mathrm{D}_2 := \mathrm{D}_{(0,1,0,0)}$, $\mathrm{D}_4 := \mathrm{D}_{(0,0,0,1)}$ and $\mathrm{D}^\circ = \mathrm{D}_{(1,1,1,1)}$. The eigenvalue λ_i of D_i for $i \in \{1, 2, 4\}$ is

$$\lambda_1 = \eta_1, \quad \lambda_2 = \frac{\eta_1 \eta_2}{\eta_1 + \eta_2}, \quad \lambda_4 = \eta_2.$$

Shorting like in (9) we see that $\infty \mathrm{D}_1 + \mathrm{D}$ has only one eigenvector (up to positive multiples) defined by the conductances $(\infty, 0, 0, 1)$ with eigenvalue η_2. For the same reason $\infty \mathrm{D}_4 + \mathrm{D}$ has only the eigenvector defined by the conductances $(1, 0, 0, \infty)$ with eigenvalue η_1. Furthermore, $\infty \mathrm{D}_2 + \mathrm{D}$ consists of eigenvectors with eigenvalue $\eta_1 + \eta_2$. Thus

$$\gamma_1 = \eta_2, \quad \gamma_2 = \eta_1 + \eta_2, \quad \gamma_4 = \eta_1.$$

The repellence conditions $\lambda_2 < \gamma_2$ of Sabot's test is true for all positive coupling weights. But $\lambda_1 < \gamma_1$ and $\lambda_4 < \gamma_4$ are contradictory. Even worse, they imply nonexistence. The only alternative left is to require $\eta_1 = \eta_2$.

In the case $\eta_1 = \eta_2 = 1$ we have an additional symmetry which allows to set $d_1 = d_4$. The new D is only three dimensional with conductances $d_1, d_2, d_3 \in \mathrm{R}_+$ and $\mathrm{D}_1 := \mathrm{D}_{(1,0,0)}$, $\mathrm{D}_2 := \mathrm{D}_{(0,1,0)}$ and $\mathrm{D}^\circ = \mathrm{D}_{(1,1,1)}$ are the only Λ-invariant D-parts. This time

$$\lambda_1 = 1, \quad \lambda_2 = \tfrac{1}{2}.$$

Shorting again we derive

$$\gamma_1 = \tfrac{1}{2}, \quad \gamma_2 = 2.$$

Thus $\gamma_1 < \lambda_1$ and we deduce again nonexistence.

References

[1] W.N. Anderson, G.E. Trapp, Shorted operators, II, SIAM J. Appl. Math. 28 (1) (1975) 60–71.

[2] W.N. Anderson, G.E. Trapp, Matrix operations induced by electrical network connections—a survey, in: C.V. Coffman, G.J. Fix (Eds.), Mathematical Models, Academic Press, New York, 1979, pp. 53–73.
[3] M.T. Barlow, Random walks, electrical resistance and nested fractals, in: K.D. Elworthy, N. Ikeda (Eds.), Asymptotic Problems in Probability Theory Stochastic Models and Diffusions on Fractals, Pitman Research Notes in Mathematics Series, vol. 283, Longman Scientific & Technical, Harlow, 1993, pp. 131–157.
[4] M.T. Barlow, R.F. Bass, Brownian motion and harmonic analysis on Sierpinski carpets, Canad. J. Math. 51 (4) (1999) 673–744.
[5] R.A. Brualdi, S.V. Parter, H. Schneider, The diagonal equivalence of a nonnegative matrix to a stochastic matrix, J. Math. Anal. Appl. 16 (1966) 31–50.
[6] P.G. Doyle, J.L. Snell, Random Walks and Electrical Networks, Carus Mathematical Monographs, vol. 22, Mathematical Association of America, Washington, DC, 1984.
[7] G.A. Edgar, Measure, Topology and Fractal Geometry, Undergraduate Texts in Mathematics, Springer, New York, 1990.
[8] M. Fukushima, Y. Oshima, M. Takeda, Dirichlet Forms and Symmetric Markov Processes, de Gruyter Studies in Mathematics, vol. 19, de Gruyter, Berlin, New York, 1994.
[9] B.M. Hambly, S.O.G. Nyberg, Finitely ramified graph directed fractals, spectral asymptotics and the multidimensional renewal theorem, Proc. Roy. Soc. Edinburgh Sect. A 46 (2002) 1–34.
[10] T. Hattori, H. Hattori, H. Watanabe, Gaussian field theories on general networks and the spectral dimension, Progr. Theoret. Phys. Suppl. 92 (1987) 108–143.
[11] H. Heuser, Lehrbuch der Analysis. Teil 1, Teubner, Stuttgart, 1980.
[12] A. Karlsson, G.A. Noskov, The Hilbert metric and Gromov hyperbolicity, Enseign. Math. II Ser. 48 (2002) 73–89.
[13] J. Kigami, Analysis on Fractals, Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001.
[14] J. Kigami, R. Strichartz, K.C. Walker, Constructing a Laplacian on the diamond fractal, Experiment. Math. 10 (2001) 437–448.
[15] U. Krause, Concave Perron–Frobenius theory and applications, Nonlinear Anal. 47 (2001) 1457–1466.
[16] M.G. Krein, M.A. Rutman, Linear operators leaving invariant a cone in a Banach space, Uspekhi Mat. Nauk 31 (23) (1948) 3–95 (English transl. Amer. Math. Soc. Transl. 26 (1950) 199–325).
[17] T. Lindstrøm, Brownian Motion on Nested Fractals, Memoirs of American Mathematical Society, vol. 83(420), American Mathematical Society, Providence, RI, 1990.
[18] J. Mallet-Paret, R.D. Nussbaum, Eigenvalues for a class of homogeneous cone maps arising from max-plus operators, Discrete Contin. Dynamic System 8 (3) (2002) 519–562.
[19] R.D. Mauldin, S.C. Williams, Hausdorff dimension in graph directed constructions, Trans. Amer. Math. Soc. 309 (1988) 811–829.
[20] V. Metz, How many diffusions exist on the Vicsek snowflake?, Acta Appl. Math. 32 (1993) 227–241.
[21] V. Metz, Hilbert's projective metric on cones of Dirichlet forms, J. Funct. Anal. 127 (1995) 438–455.
[22] V. Metz, Shorted operators: an application in potential theory, Linear Algebra Appl. 264 (1997) 439–455.
[23] V. Metz, The cone of diffusions on finitely ramified fractals, Nonlinear Anal. 55 (2003) 723–738.
[24] V. Metz, The Laplacian of the Hany fractal, Arab. J. Sci. Eng. Sect. C Theme Issues 28 (1C) (2003) 199–211.
[25] V. Metz, A note on the diamond fractal, Potential Anal. 21 (2004) 35–46.
[26] H. Minc, Nonnegative Matrices, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, New York, 1988.
[27] R.D. Nussbaum, Convexity and log convexity for the spectral radius, Linear Algebra Appl. 73 (1986) 59–122.
[28] R.D. Nussbaum, Hilbert's Projective Metric and Iterated Nonlinear Maps, Memoirs of American Mathematical Society, vol. 75(391), American Mathematical Society, Providence, RI, 1988.

[29] R.D. Nussbaum, Iterated Nonlinear Maps and Hilbert's Projective Metric, II, Memoirs of American Mathematical Society, vol. 79(401), American Mathematical Society, Providence, RI, 1989.

[30] R. Peirone, Convergence and uniqueness problems for Dirichlet forms on fractals, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. 3-B (8) (2000) 431–460.

[31] C. Preston, Iterates of Maps on an Interval, Lecture Notes in Mathematics, vol. 999, Springer, Berlin, Heidelberg, 1983.

[32] C. Sabot, Existence and uniqueness of diffusions on finitely ramified self-similar fractals, Ann. Sci. Ecole Norm. Sup. 30 (4) (1997) 605–673.

[33] E. Seneta, Non-Negative Matrices. An Introduction to Theory and Applications, George Allen & Unwin, 1973.

[34] D.R. Smart, Fixed Point Theorems, Cambridge Tracts Mathematics, Cambridge University Press, Cambridge, 1974.

[35] R.J. Webster, Convexity, Oxford Science Publications, Oxford University Press, Oxford, 1944.